A TEST OF THE TEMPORAL STABILITY OF PROPORTIONAL HAZARDS MODELS FOR PREDICTING BANK FAILURE

Kathleen L. Henebry

Abstract

This research uses both cash flow and non-cash flow proportional hazards models to test for stability of the models over time. Several different time horizons and start dates were used to test stability over the 1985-1989 time period. The results indicate that none of the specific formulations were stable across different starting dates nor across different horizons for the same starting date. Forecast models further tested stability and only three variables were found to be consistently useful in predicting bank failure: Primary Capital to Total Assets (PCTA), Nonperforming Loans to Total Loans (NPLTL) and Total Loans to Total Assets (TLTA).

INTRODUCTION

During the 1980s the rate and costs of bank and savings and loan failures soared. It is important to learn as much as possible about why these failures happened and how to predict such failures before more of them occur in the future, creating additional large costs and disruptions to depositors and the public.

As bank managers and regulators attempt to improve their strategic decision methods and models to prevent future failures more attention should be given to cash flow. To date, very little has been published on the impact of including cash flow variables upon bank failure prediction models. This paper will extend the work of a few earlier papers on that issue in an attempt to further define the role such variables should take in future plans and decisions by management and regulators.

This paper extends the research of Henebry [3] which focused on the addition of cash flow variables to a Cox Proportional Hazards (PHM) model for bank failure prediction. The models used in Henebry [3] also contained regular accounting ratios representing the CAMEL rating system areas of capital, asset quality, earnings and liquidity. This paper will utilize the same model structures.

The focus of this paper is to extend the time frames examined and determine if the previous models are stable across different start dates. The Henebry [3] paper estimated sets of models with and without the cash flow variables with five-, four-, three-, two- and one-year time frames all starting in 1985. This paper will estimate four-, three-, two- and one-year models with start dates of 1985, 1986, 1987, 1988 and 1989. Given that the upper limit on the data set is an end point of 1990, not every time frame can be considered for every start date.

The 1985 and 1986 data will be used to generate models of one, two, three and four years. As the starting point is moved forward the length of the longest model is shortened to three, then two and finally one year.

The modeling technique used in this paper is the same used in Lane, Looney and Wansley [5], Whalen [11] and Henebry [3]. These papers use models which generate information on both probability of failure and probable time to failure; a proportional hazards model (PHM), specifically the version developed by Cox [1], which uses an exponential hazard function. This model assumes that whatever the effect of the variables on survival (or failure), it will be proportional across time. Readers wishing for mathematical details of the modeling technique are referred to Appendix A and these earlier papers.

^{*}University of Nebraska at Omaha

This paper is based on my dissertation. I would like to acknowledge the helpful comments of my chair and members of my committee as well as participants at the Eastern Finance Association Annual Meeting of 1996.

DATA AND METHODOLOGY

The data set used in this paper is derived from FDIC annual reports which identified failed banks. The failed banks then were matched with non-failed banks on the bases of size and geographic proximity to create the data set to which the model is applied. The data set was split in half; one half for estimating the model and the other to be used for out of sample testing. The cash flow calculations are the same as in Henebry [3]. A complete list of the variables used is shown in Table 1.

Category	Description	Code Name
Cash Flow:	net cash flow after operating expenses total flow	RNOCF**
	change in income earned but not collected total flow	RCIEBNC
	net investment cash flow total flow	RNICF ^{**}
	net change in sources of funds total flow	RNCSF ^{**}
	net interest paid on non - deposit sources of funds total flow	RNINDS
	net other assets and liabilities total flow	RNOAL
	dividends total flow	RDIV
	total cash flow total assets	RFLOWTA
Capital:	primary capital average total assets	РСТА
	net PCTA = PCTA - (total nonperforming loans ÷ average total assets)	NPCTA
Asset Quality:	total nonperforming loans total loans	NPLTL
	net charge offs net loans	NCONL
Management:	commercial real estate construction loans total assets	RECLTA
	commercial and industrial loans total assets	CILTA
Earnings:	operating expenses average total assets	OETA
	return on assets	ROA
Liquidity:	total loans total assets	TLTA
	total domestic deposit denominations of \$100,000 or more total assets	CD100TA

TABLE 1 Variable List

Category	Description	Code Name
Disaggregated Cash Flow:	change in loans total flow	CHLNS
	interest income total flow	RINTINC
	interest expense total flow	RINTEXP
	salary and benefit expense total flow	RSALEXP
	change in deposits total flow	RCHDEP
	change in total brokered deposits total flow	RBRKTOT
	change in brokered deposits of denominations less than \$100,000 total flow	RBRKSML
	change in brokered deposits of denominations greater than \$100,000 but brokered in denominations less than \$100,000 total flow	RBRKLG
	change in non - transaction savings accounts total flow	RSAVACT
	CDs denominations less than \$100,000 total flow	RCDSML
	change in CDs denominations greater than \$100,000 total flow	CHCD100

TABLE 1 (CONT'D) Variable List

**Not included when disaggregated cash flow variables are used

Several models were estimated with different survival horizons, specifically, horizons of one to four years. For example, data from 1985 is used in a four-year model to predict probable failures occurring any time between 1986 and 1989. Similarly, 1985 data is used in a three-year model for probable failures between 1986 and 1988. Three-, two- and one-year models were also estimated for the 1985 data. The different models used different data sets i.e. different banks will be censored (fail) at the different horizons.

This process was repeated moving forward in time to generate several models for each year of starting data. This resulted in two four-year models using 1985 and 1986 data, three three-year models using 1985, 1986 and 1987 data, four two-year models using 1985, 1986, 1987 and 1988 data and five one-year models using 1985, 1986, 1987, 1988 and 1989 data.

This process allows a study of the temporal stability of the models. If the same variables remain significant over time, the models will not need to be re-estimated every year to remain useful. It also allows a determination of the stability of the models across different time frames for the same starting point.

One final stability test was also performed; using the significant variables from the 1985 two-year model to predict failures in the 1986, 1987 and 1988 two-year data sets.

RESULTS

General Temporal Stability

For clarity of discussion the models will be referred to by a three-digit number denoting the start date and the horizon length; 854 being the four-year models starting in 1985. The RA designation indicates the non-cash flow models while the RCA designation indicates the cash flow models.

An examination of the results for the four-year models from Table 2 shows several notable differences. Model 854RCA has nine significant variables including Changes in Income Earned But Not Collected (RCIEBNC), Net Other Assets and Liabilities (RNOAL), Commercial and Industrial Loans (CILTA) and Real Estate and Construction Loans (RECLTA), none of which appear in the smaller set (six) of significant variables for 864RCA.

These results indicate that commercial loans may be an unstable indicator of bank failure, more highly subject to the business cycle than other variables such as Primary Capital to Total Assets (PCTA), Total Loans to Total Assets (TLTA) and Nonperforming Loans to Total Loans (NPLTL) which appear in both cash flow and non-cash flow models for the four-year horizon.

The non-cash flow models (854RA and 864RA) appear more stable at the four-year horizon, with three of the five significant variables in the 854 model being retained for the 864 model. The RECLTA variable again only appears in the 854 model, reinforcing the comments above with respect to the instability of commercial loans as a failure indicator.

The three-year cash flow models, 853RCA, 863RCA and 873RCA have only three variables which are stable across all three time frames; TLTA, PCTA and NPLTL. There is a wide variation in the other significant variables of these three models. Model 853RCA has nine significant variables, of which five are cash flow variables, while model 863RCA has no significant cash flow variables at all.

The non-cash flow models are somewhat more stable, but the only three variables which carry through all models are TLTA, PCTA and NPLTL. All three of these variables are commonly used by regulatory agencies in assessing bank performance. Further examination of Table 2 shows that these three variables are the most stable across different time horizon lengths and different starting dates. Not surprisingly, PCTA appears as a significant variable in every model, and NPLTL appears in all but four models. The TLTA variable is somewhat less significant, appearing in only 12 of the 30 models.

Only in the four-year models do any of the cash flow variables track forward to a later starting date; Net Investment Cash Flow (RNICF) and Net Changes in Sources of Funds (RNCSF) appear in both 854RCA and 864RCA. No other time horizons have cash flow variables which carry through from the 1985 start date models. These results indicate that neither the cash flow nor non-cash flow models are stable over time.

It should also be noted that the cash flow variable, RNICF, is fairly strongly correlated with the regular accounting ratios of PCTA and Net Chargeoffs to Net Loans (NCONL). These correlations may be a contributing factor to RNICF appearing in some, but not all, of the models as a significant variable. In every case in which RNICF is significant PCTA is significant and in two cases NCONL is also significant.

Further examination of Table 2 reveals that there is a lack of stability across different time horizons even within the same start date. For example, there is little overlap in significant variables for either the cash flow or non-cash flow 854, 853, 852 and 851 models. Similar results are found for all other start date model sets.

One possible explanation for the temporal instability across models with the same start point, but different horizons, may lie in another of the underlying assumptions of the PHM itself, namely that the variables are assumed to be unchanging over the entire horizon. It is very likely that a variable which is not significant over a one-year time horizon may become significant over a longer horizon.

For example, one year of poor operating cash flow, poor investment cash flow or rising costs of funds may not be significant in predicting failure whereas two or more consecutive years of such results are significant. This may explain the more frequent appearance of the cash flow variables as significant in models with longer time horizons.

The differences in models across different start times may be largely explained by changes in the values of the variables over time. The values of the explanatory variables do not remain constant from 1985 through 1989 as confirmed by changes in the mean values across the different data sets. As the variables' values change their relative importance in predicting failure may also vary depending on the size and direction of the changes for a specific variable.

Significant Variable	Model
RCIEBNC RNOAL TLTA CILTA PCTA	851RCA
TLTA CILTA PCTA NPLTL	851RA
RCIEBNC RNICF RNCSF TLTA PCTA NPLTL	852RCA
TLTA RECLTA PCTA NCONL NPLTL	852RA
RNOCF RNICF RNCSF RNOAL RDIV TLTA RECLTA PCTA NPLTL	853RCA
TLTA RECLTA PCTA NCONL NPLTL	853RA
RCIEBNC RNICF RNCSF RNOAL TLTA CILTA RECLTA PCTA NPLTL	854RCA
TLTA RECLTA PCTA NCONL NPLTL	854RA

TABLE 2 Significant Variables

Significant VariablesModelsRNOCF RNICF RNOAL TLTA PCTA NCONL NPLTL855RATLTA RECLTA PCTA NCONL NPLTL855RARNOCF RNOCF CDI00TA PCTA NPLTL861RCARNOCF CDI00TA PCTA NPLTL861RATLTA PCTA NOCN NPLTL861RATLTA PCTA NOCTA NPLTL861RARNOCF CDI00TA PCTA NPLTL861RATLTA CD100TA862RCATLTA CD100TA PCTA NPLTL862RATLTA CD100TA PCTA NPLTL863RACD100TA PCTA NPLTL863RACD100TA PCTA NPLTL863RACD100TA PCTA NPLTL863RACD100TA PCTA NPLTL863RACD100TA PCTA NPLTL864RARNOCF RNCSF TLTA PCTA NPLTL864RARNOCF RNCSF TLTA PCTA NPLTL864RA		
RNICF RNCSF RNOAL TLTA PCTA NCONL NPLTL TLTA RECLTA PCTA NCONL NPLTL RNOCF RNO	Significant Variables	Models
RECLTA PCTA NCONL NPLTL RNOCF CD100TA PCTA NPLTL TLTA CD100TA PCTA NPLTL CD100TA PCTA NPLTL CD100TA PCTA NPLTL TLTA CD100TA PCTA NPLTL TLTA CD100TA PCTA NPLTL TLTA CD100TA PCTA NPLTL TLTA CD100TA PCTA NPLTL TLTA CD100TA PCTA NPLTL TLTA CD100TA PCTA NPLTL TLTA CD100TA PCTA NPLTL TLTA CD100TA PCTA NPLTL TLTA CD100TA PCTA NPLTL TLTA CD100TA PCTA NPLTL TLTA CD100TA PCTA NPLTL TLTA CD100TA PCTA NPLTL TLTA CD100TA PCTA NPLTL TLTA CD100TA PCTA NPLTL TLTA CD100TA PCTA NPLTL TLTA CD100TA PCTA NPLTL TLTA CD100TA PCTA NPLTL TLTA S63RA CD100TA PCTA NPLTL RNOCF RNO	RNICF RNCSF RNOAL TLTA PCTA NCONL	855RCA
RNICF CD100TA PCTA NPLTL TLTA ROCF CD100TA RNOCF TLTA CD100TA PCTA NPLTL TLTA ROCF TLTA ROCF	RECLTA PCTA NCONL	855RA
PCTA NPLTL CD100TA RNOCF TLTA CD100TA PCTA NPLTL TLTA CD100TA PCTA NPLTL TLTA CD100TA PCTA NPLTL TLTA CD100TA PCTA NPLTL TLTA CD100TA PCTA NPLTL TLTA CD100TA PCTA NPLTL TLTA CD100TA PCTA NPLTL TLTA CD100TA PCTA NPLTL TLTA CD100TA PCTA NPLTL TLTA RNOCF RNCSF TLTA NPLTL TLTA RNOCF RNCSF TLTA NPLTL TLTA RNCSF TLTA NPLTL RNOCF RNCSF TLTA NPLTL TLTA RNCSF TLTA NPLTL TLTA RNCSF TLTA NPLTL	RNICF CD100TA PCTA	861RCA
TLTA CD100TA PCTA NPLTL TLTA CD100TA PCTA NPLTL TLTA CD100TA PCTA NPLTL TLTA CD100TA PCTA NPLTL TLTA R63RA CD100TA PCTA NPLTL TLTA R0CF R0C	PCTA NPLTL	861RA
CD100TA PCTA NPLTL TLTA 863RCA CD100TA PCTA NPLTL TLTA 863RA CD100TA PCTA NPLTL RNOCF 864RCA RNICF RNCSF TLTA PCTA NPLTL TLTA 864RA ROA PCTA	TLTA CD100TA PCTA	862RCA
CD100TA PCTA NPLTL TLTA 863RA CD100TA PCTA NPLTL RNOCF 864RCA RNICF RNCSF TLTA PCTA NPLTL TLTA 864RA ROA PCTA	CD100TA PCTA	862RA
CD100TA PCTA NPLTL RNOCF 864RCA RNICF RNCSF TLTA PCTA NPLTL TLTA 864RA ROA PCTA	CD100TA PCTA	863RCA
RNICF RNCSF TLTA PCTA NPLTL TLTA 864RA ROA PCTA	CD100TA PCTA	863RA
ROA PCTA	RNICF RNCSF TLTA PCTA	864RCA
	ROA PCTA	864RA

Significant Variables	Models
RNCSF RNOAL CD100TA PCTA NCONL	871RCA
CD100TA ROA PCTA NCONL NPLTL	871RA
CD100TA PCTA NPLTL	872RCA
TLTA CD100TA PCTA NPLTL	872RA
RNICF RNCSF RNOAL CD100TA PCTA NCONL NPLTL ROA	873RCA
TLTA CD100TA ROA PCTA NCONL NPLTL	873RA
CILTA CD100TA PCTA NPLTL	881RCA
CILTA CD100TA PCTA NPLTL	881RA
RCIEBNC CILTA PCTA NPLTL	882RCA
CILTA PCTA NPLTL	882RA
CILTA PCTA	891RCA
CILTA PCTA	891RA

Forecast Model Stability

It is important to determine whether or not the significant variables of one model will remain significant when applied to data sets outside the time frame used to estimate the model. If such stability can be demonstrated, the models will be far more useful for forecasting purposes in predicting future failures. To this end, the author ran the 862, 872 and 882 data sets against the significant variables found using the 852 data set.

The 852RCA model's significant variables were used to run 852/62RCA, 852/72RCA and 852/82RCA models. The initial significant variables from the 852RCA model were RCIEBNC, RNICF, RNCSF, TLTA, PCTA, NPLTL. Of these only TLTA, PCTA and NPLTL remained significant in the forecast models. These three variables were significant in all three forecast models, although for the 852/82RCA model TLTA was significant only at the 6% level rather than the 5% level. None of the cash flow variables which were significant in the 852RCA model, RCIEBNC, RNICF, RNCSF proved significant for the later data sets. These results are summarized in Table 3.

Significant Variable ^a	Model
TLTA PCTA NPLTL	852/62RCA
TLTA PCTA NPLTL	852/62RA
TLTA PCTA NPLTL	852/72RCA
TLTA PCTA NPLTL	852/72RA
TLTA* PCTA NPLTL	852/82RCA
TLTA PCTA NPLTL	852/82RA

TABLE 3 Significant Variables for Forecast Models Using 852 Significant Variables with 862, 872 and 882 Data Sets.

a. Variable significant at 5% unless otherwise specified.

*TLTA for model 852/82RCA is significant at the 6% level.

Table 3 also shows that for the 852/62RA, 852/72RA and 852/82RA models the only three significant variables at the 5% level are again, TLTA, PCTA and NPLTL. These results indicate that models containing variables other than these three will be unstable across time. The results also indicate that these three variables are the most consistently useful in predicting bank failure and dominate other variables placed in the models over time.

From the (-2logL) values in Table 4 it is clear that none of the three cash flow forecast models is significantly different from the non-cash flow forecast models. None of the differences in (-2logL) exceed the 5% significance critical value of 3.84.

An analysis of the Type I and Type II error rates for the forecast models is summarized in Table 5. This table shows that the Type I error rates are very small, actually zero for two of the three forecast models, while the Type II error rates are higher than those found in earlier papers.

Earlier papers such as Martin [6], Pettway and Sinkey [7], Richardson and Davidson [8], Rose and Kolari [9], Spahr [10], Jordan and Henderson [4] and Espahbodi [2] which all used a multiple discriminant analysis (MDA) technique all had basically similar Type II error rates. These error rates ranged from 5% to just under 20% incorrect classification of nonfailed banks.

Whalen [11] which ran models for various time horizons found Type II error rates ranging from 1.8% to 14.5%. The 14.5% figure consisted of 238 banks, 92 of which failed after the time horizon cutoff for the study. A Type II error which fails shortly after the end of the study is an example of the model correctly predicting failure, but at an early date.

TABLE 4Log Likelihood Ratios, and Chi-square Statistics forForecast Models Using 852 Significant Variableswith 862, 872 and 882 Data Sets.

Model Final Versions	-2logL	Chi-square Statistic ^a
852/62RCA	1410.428	
852/62RA	1406.988	3.44
852/72RCA	1122.670	
852/72RA	1124.893	2.22
852/82RCA	1007.160	
852/82RA	1008.648	1.49

a. chi-square statistic for differences in -2logL

*none are significant at the 5% level

TABLE 5Prediction Results with Type I and Type II ErrorRates for Forecast Models Using 852 SignificantVariables with 862, 872 and 882 Data Sets.

NAME	TYPE 1	CPF	CPNF	TYPE 2	TOTAL
852/62RA	2 1.27%	156 98.73%	67 42.41%	91 57.59%	316
852/62RCA	2 1.27%	156 98.73%	69 43.67%	89 56.33%	316
852/72RA	0 0.00%	130 100.00%	29 21.01%	109 78.99%	268
852/72RCA	0 0.00%	130 100.00%	31 22.46%	107 77.54%	268
852/82RA	0 0.00%	115 100.00%	30 24.59%	92 75.41%	237
852/82RCA	0 0.00%	115 100.00%	36 29.51%	86 70.49%	237

Table 6 summarizes the results for McNemar's test for significance of changes for the forecast models. The results indicate that only the 852/82 models yield significantly different predictions when comparing the cash flow model versus the non-cash flow model. This difference must be driven by the slight difference in significance for the TLTA variable as mentioned above, for there are no other differences in the significant variables between the two models.

TABLE 6
McNemar's Test for Significance of Changes
G-Statistic for Forecast Models Using 852 Significant
Variables with 862, 872 and 882 Data Sets.

Year	0,0	0,1	1,0	1,1	G
852/62	245	2	0	69	2.7682
852/72	235	4	2	27	0.6787
852/82	200	7	1	29	5.0593*

*Significant at 5%

Notes:

1) 0,0: Both models predict failure. Corresponds with 'a' in McNemar's formula.

2) 0,1: RA model predicts failure. RCA model predicts survival. Corresponds with 'b' in McNemar's formula.

3) 1,0: RA model predicts survival. RCA model predicts failure. Corresponds with 'c' in McNemar's formula.

4) 1,1: Both models predict survival. Corresponds with 'd' in McNemar's formula.

The Spearman rank order correlation coefficients shown in Table 7 for all of the forecast models are significant; all exceed 0.90. The critical value for rejecting the hypothesis of no correlation is no larger than 0.195 at the 5% level or 0.254 at the 1% level for even the smallest sample size used; the calculated coefficient values indicate very strong correlation between the predicted time to failure and the actual time to failure for all of the forecast models; both cash flow and non-cash flow versions.

TABLE 7

Spearman's Rank Order Correlation Coefficients: Correlation Between Actual Failure Rank and Modeled Failure Rank for Forecast Models Using 852 Significant Variables with 862, 872 and 882 Data Sets.

Model	r-value
852/62RA	0.939
852/62RCA	0.939
852/72RA	0.930
852/72RCA	0.930
852/82RA	0.930
852/82RCA	0.931

SUMMARY AND FURTHER RESEARCH

No temporally stable models were found using either the cash flow or the non-cash flow variable sets. Only three variables appear in the majority of models; PCTA, TLTA and NPLTL. Even these three variables are not found together in all of the tested models.

For the forecast models using a single set of variables found to be significant for the two-year model starting with year end 1985 data applied to the two-year data sets starting with year end 1986, year end 1987 and year end 1988 data no cash flow variables remained significant. This indicates that the addition of the cash flow variables did not improve the performance of the models. (See Tables 3-7.)

At this point it might appear that bank managers can safely ignore cash flow when developing their strategic plans. But, this runs counter to standard thought in finance and business management for nonfinancial firms. It should be noted, however, that the definition of failure here is the decision by a regulatory agency to close or merge the bank. These decisions were made on models which, so far as can be determined by an outside observer, ignore cash flow variables. It would be of great interest to see what the results would be under conditions where the regulatory closure decision also considered cash flow.

Although including cash flow variables, as they were formulated in this study, does not improve the stability of the Cox PHMs there is still work to be done in the area of the relationship between cash flow and bank failure. The Cox models as estimated here do not allow the variables to change over the time frame of the study. A modification of the PHM which allows for changes in the predictive variables over time may yield different results.

It is possible that the volatility of cash flow is more important than its actual value at any point in time. It may prove useful to estimate a PHM using measures of the variance in the cash flow variables used in this study, rather than their year end values. Using a volatility measure may remove some of the instability observed in the cash flow variables used in this paper.

In summation, while this particular formulation has not yielded positive support for a stable relationship between either cash flow and bank failure, or most non-cash flow ratios used in previous studies and bank failure, there are still avenues of research to explore in this area. The failure of the current cash flow models to more accurately predict future bank failures should not be considered sufficient evidence for bank managers or regulators to ignore cash flow in strategic planning.

REFERENCES

- [1] Cox, D.R., "Regression Models and Life-Tables," Journal of the Royal Statistical Society, Series B, 1972, pp. 187-220.
- [2] Espahbodi, P., "Identification of Problem Banks and Binary Choice Models," *Journal of Banking and Finance*, February 1991, pp. 53-71.
- [3] Henebry, K.L, "Do Cash Flow Variables Improve the Predictive Accuracy of a Cox Proportional Hazards Model for Bank Failure?" *Quarterly Review of Economics and Finance* 36:3, 1996.
- Jordan, C.E. and J.R. Henderson, "Evaluating the Financial Health of Midwestern Banks," *Ohio CPA Journal* 49:2, 1990, pp. 42-46.
- [5] Lane, W.R., S.W. Looney and J.W. Wansley, "An Application of the Cox Proportional Hazards Model to Bank Failure," *Journal of Banking and Finance*, 1986, pp. 511-531.
- [6] Martin, D., "Early Warning of Bank Failure: A Logit Regression Approach," *Journal of Banking and Finance* 1, 1977, pp. 249-276.
- [7] Pettway, R.H., and J. Sinkey, Jr., "Establishing On-site Bank Examination Priorities: An Early Warning System Using Accounting and Market Information," *Journal of Finance*, March 1980, pp. 137-150.
- [8] Richardson, F.M., and L.F. Davidson, "An Exploration Into Bankruptcy Discriminant Model Sensitivity," *Journal of Business Finance & Accounting* 10:2, 1983, pp. 195-207.
- [9] Rose, P.S., and J.W. Kolari, "Early Warning Systems as a Monitoring Device for Bank Condition," *Quarterly Journal of Business* and Economics 24:1, 1985, pp. 43-59.
- [10] Spahr, R.W., "Predicting Bank Failures and Intertemporal Assessment of Bank Risk," *Journal of Business Research* 19, 1989, pp. 179-185.
- [11] Whalen, G., "A Proportional Hazards Model of Bank Failure: An Examination of its Usefulness as an Early Warning Tool," *Economic Review, Federal Reserve Bank of Cleveland* 27:1, 1991, pp. 21-31.

APPENDIX A Cox Proportional Hazards Model

For estimating the time to failure, T, of a bank define a survivor function to be the probability that the bank survives past t time units as:

Equation 1

S(t) = Pr[T > t]

The dependent variable is time to failure, *t*.

The most common characterization of the distribution for the time to failure, t, is in terms of the hazard function:

Equation 2

$$h(t) = \lim_{dt \to 0} \frac{P(t < T < t + dt | T > t)}{dt} = \frac{-S'(t)}{S(t)}$$

This function represents the probability of failure in the next instant assuming the bank has survived until *t*. The Cox model is then given by:

Equation 3

$$h(t \mid X, B) = h_0(t)exp(X'B)$$

where X and B (beta) are vectors of variants and regression coefficients. The variants are, of course, assumed to affect the probability of failure and the coefficients are the model's estimates of how they do so.

No distributional assumptions are needed for $h_0(t)$, the hazard function, or for the estimation of the coefficients. The baseline hazard function in the Cox model is completely arbitrary; no prior assessment of the hazard function is needed before the model is estimated.

The estimated coefficient vector depends only on the rank order of the dependent variable and is invariant with respect to monotonic transformations of the dependent variable.

Type of Cash Flow	Inputs
Operating:	
Inflows	Interest Income + Non-interest Income +
Outflows	↓ Trade Account Salary & Benefit Expense + Equipment Expense + Interest Paid on Deposits + Other Operating Expense + Taxes + ↑ Trade Account
Net Operating Cash Flows (NOCF)	Inflows - Outflows
Investment:	
Inflows	↓ [Loans - Net Chargeoffs] + ↓ Securities + ↓ Plant & Equipment + ↓ Other Real Estate Owned + ↓ FFS & Repos
Outflows	 ↑ [Loans - Net Chargeoffs] + ↑ Securities + ↑ Plant & Equipment + ↑ Other Real Estate Owned + ↑ FFS & Repos
Net Investment Cash Flows (NICF)	Inflows - Outflows
Receivables:	
Inflows Outflows Change in Accounts Receivable is change in income earned but not collected (CIEBNC)	↓ Income Earned But Not Collected ↑ Income Earned But Not Collected Inflows - Outflows
Payables:	
Inflows Outflows Change in Accounts Payable (CAP) Change in Other Current Assets (COCA) Change in Other Current Liabilities (COCL)	 ↑ Expenses Payable ↓ Expenses Payable Inflows - Outflows ↓ OCA - ↑ OCA ↑ OCL - ↓ OCL
Financing:	
Inflows Outflows	 ↑ Subordinated Debt + ↑ FFP & Reverse Repos + ↑ Treasury Demand Notes + ↑ Other Borrowings + ↑ Common Stock + ↑ Preferred Stock + ↑ Mortgages + ↑ Deposits ↓ Subordinated Debt + ↓ FFP & Reverse Repos + ↓ Treasury Demand Notes + ↓ Other Borrowings + ↓ Common Stock + ↓ Preferred Stock + ↓ Mortgages + ↓ Deposits
Net Change in Sources of Funds (NCSF) Net Interest Paid on Nondeposit Sources of Funds (NINDS) Dividends (DIV)	Inflows - Outflows Sum of Interest Paid on any source of funds which is not a deposit of any type Dividend Declared on Common Stock + Dividend Declared on Preferred Stock
Net Other Assets & Liabilities (NOAL) Change in Cash (CHCASH)	Value such that cash flows will balance with change in cash [Ending Balance Cash & Due From] - [Beginning Balance Cash & Due From]

APPENDIX B Cash Flow Calculations