

WP/05/219

IMF Working Paper

Estimating Markov Transition Matrices Using Proportions Data: An Application to Credit Risk

Matthew T. Jones

IMF Working Paper

Monetary and Financial Systems Department

**Estimating Markov Transition Matrices Using Proportions Data:
An Application to Credit Risk**

Prepared by Matthew T. Jones¹

Authorized for distribution by S. Kal Wajid

November 2005

Abstract

This Working Paper should not be reported as representing the views of the IMF.

The views expressed in this Working Paper are those of the author(s) and do not necessarily represent those of the IMF or IMF policy. Working Papers describe research in progress by the author(s) and are published to elicit comments and to further debate.

This paper outlines a way to estimate transition matrices for use in credit risk modeling with a decades-old methodology that uses aggregate proportions data. This methodology is ideal for credit-risk applications where there is a paucity of data on changes in credit quality, especially at an aggregate level. Using a generalized least squares variant of the methodology, this paper provides estimates of transition matrices for the United States using both nonperforming loan data and interest coverage data. The methodology can be employed to condition the matrices on economic fundamentals and provide separate transition matrices for expansions and contractions, for example. The transition matrices can also be used as an input into other credit-risk models that use transition matrices as a basic building block.

JEL Classification Numbers: C13

Keywords: Markov transition matrix; credit risk; nonperforming loans; interest coverage

Author(s) E-Mail Address: mjones2@imf.org

¹ The author is grateful to colleagues in MFD for useful comments on earlier drafts of this paper, and to Marianne El-Khoury and the FDIC for assistance with data.

Contents	Page
I. Introduction	3
II. Credit Quality Dynamics Using Transition Matrices.....	4
A. Transition Matrices When Individual Transitions Known	4
B. Transition Matrices When Individual Transitions Unknown.....	6
III. Application to U.S. Data.....	8
A. Nonperforming Loan Data	8
B. Interest Coverage Data	12
IV. Additional Applications.....	15
V. Conclusions.....	16
Technical Appendix I.....	17
A. Estimating Transition Probabilities with Observed Transitions	17
B. Markov Probability Model.....	18
C. Data Descriptions	22
References.....	24
Tables	
1. Estimated Quarterly Transition Matrix for U.S. Commercial Bank Loans and Leases	10
2. Estimated Quarterly Transition Matrix, Split Sample	11
3. Estimated Quarterly Transition Matrix, Low and High Growth Estimates	12
4. Estimated Annual Transition Matrix Using Interest Coverage Ratio	14
Figures	
1. U.S. Nonperforming Loan Ratios	10
2. Real GDP Growth and Loan Ratios.....	12
3. Using Interest Coverage Data to Estimate Transition Matrices.....	13
4. Interest Coverage Ratio for U.S. Companies.....	14
5. Interest Coverage Ratio, Performing Loans, and Real GDP.....	15
Appendix Tables	
1. Illustrative Example of Using Count Proportions to Estimate Transition Probabilities.....	18

I. INTRODUCTION

The experience with banking crises in numerous countries has demonstrated the intricate links between deteriorations in creditor quality, macroeconomic conditions, and institutional failure.² The costly lessons learned from recent banking crises have illustrated the importance of proper credit-risk management to maintaining financial stability. Understanding the evolution of credit risk is thus an important step in preventing institutional failure and financial crises.

In the past 10 years there has been a dramatic increase in the analysis and understanding of the evolution of market risk, but progress in understanding credit risk has been much slower.³ Modeling credit risk is inherently more complex than modeling market risk, because the returns on a credit portfolio tend to be asymmetric, causing the distribution of returns to be highly skewed with fat negative tails. In contrast, market returns tend to be distributed more symmetrically and hence are more tractable analytically. Credit-risk events are also much less frequent than changes in market returns, and tend to be monitored less effectively, giving rise to a paucity of data. Despite these difficulties, there have been significant advances in recent years in the theory and application of credit-risk models.

One strand of the credit-risk-modeling literature makes use of a matrix of transition probabilities to explain the migration of creditor quality, as measured by proxies such as bond ratings. These models of ratings migration show the evolution of creditor quality for broad groups of creditors with the same approximate likelihood of default. This approach provides matrices of transition probabilities that can be used as an input to models of credit evolution, because they summarize a broad range of possible creditor dynamics in a simple and coherent fashion.

This paper demonstrates how to use proportions data to estimate transition matrices in circumstances where individual transitions are not observed. The paper demonstrates the application of the technique using ratio data on nonperforming loans and corporate sector interest coverage to arrive at two independent estimates of transition matrices. These estimates provide a basis for comparing official sector estimates of credit quality (derived from supervisory data) with corporate sector information on company earnings (derived from balance sheet data). The transition matrices can then be conditioned on macroeconomic variables to illustrate the impact of economic performance on creditor quality.

Because of the minimal data requirements necessary to implement the techniques shown in the paper, the approach is potentially applicable to a broad range of countries and circumstances. The methodology demonstrated in this paper can be applied to individual countries with sufficiently good data, or use average measures based on cross-country

² See Lindgren, Garcia, and Saal (1996); and Caprio and Klingebiel (1996) for surveys.

³ See Altman and Saunders (1997) for an overview of credit risk, or www.defaultrisk.com for recent papers.

experience for similar economies or financial systems. The framework can be applied prospectively for stress testing of vulnerabilities to macroeconomic conditions, or retrospectively to understand the dynamics of linkages between loan portfolios and macroeconomic outcomes.

The paper is organized as follows: Section II discusses the ratings migration literature and presents the analytic foundations of the use of proportions data to estimate transition matrices; Section III discusses the application of the methodology to nonperforming loan and corporate sector data, using information from the United States to estimate transition matrices; Section IV discusses alternative applications of the estimated transition matrices; and Section V concludes.

II. CREDIT QUALITY DYNAMICS USING TRANSITION MATRICES

A. Transition Matrices When Individual Transitions Known

In the credit-ratings literature, transition matrices are widely used to explain the dynamics of changes in credit quality. These matrices provide a succinct way of describing the evolution of credit ratings, based on a Markov transition probability model. The Markov transition probability model begins with a set of discrete credit quality ranges (or states), into which all observations (e.g., firms or institutions) can be classified. Suppose there are R discrete categories into which all observations can be ordered. We can define a transition matrix, $P = [p_{ij}]$, as a matrix of probabilities showing the likelihood of credit quality staying unchanged or moving to any of the other $R-1$ categories over a given time horizon. Each element of the matrix, p_{ij} , shows the probability of credit quality being equal to i in period $t-1$ and credit quality equal to j in period t :

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1R} \\ p_{21} & p_{22} & \cdots & p_{2R} \\ \vdots & \vdots & \ddots & \vdots \\ p_{R1} & p_{R2} & \cdots & p_{RR} \end{bmatrix}. \quad (1)$$

We impose a simple Markov structure on the transition probabilities, and restrict our attention to first-order stationary Markov processes, for simplicity.⁴ The final state, R , which can be used to denote the loss category, can be defined as an absorbing state. This means that once an asset is classified as lost, it can never be reclassified as anything else.⁵

⁴ A Markov process is stationary if $p_{ij}(t) = p_{ij}$, i.e., if the individual probabilities do not change with time. See Appendix I for more details.

⁵ Thus the final row of the transition matrix $[p_{R1} \ p_{R2} \ \dots \ p_{RR}]$ consists of zero entries everywhere, except for a one for p_{RR} on the diagonal.

Under this framework, the only relevant information for explaining the behavior of the series is its behavior in the previous period. This assumption of a first-order Markov process for credit transitions may be somewhat restrictive if credit quality responds slowly to changes in economic fundamentals, for example. Under these circumstances, using a higher-order Markov process or a longer time horizon may be more appropriate. However, using higher-order processes or longer horizons increases the complexity and data requirements quite substantially, and may not be feasible with only a limited time series. It may also be the case that credit quality itself responds quickly to changes in fundamentals, but observations on credit quality are only made infrequently. Similarly, when using some sources of information on credit quality such as supervisory data, the observed variable is not true credit quality but the supervisor's assessment of the data reported to it. Ideally, one could use hidden Markov chains to model the latent credit quality variable, using supervisory observations as the observed (or emitted) model. However, the data requirements of this approach are immense and thus are not practical for the applications considered in this paper.

Estimating a transition matrix is a relatively straightforward process, if we can observe the sequence of states for each individual unit of observation, i.e., if the individual transitions are observed. For example, if we observe the credit ratings of a group of firms at the beginning of a year and then again at the end of the year, then we can estimate the probability of moving from one credit rating to another.⁶ The probability of a firm having a particular credit rating at the end of the year, (e.g., A) given their rating at the beginning of the year (e.g., B) is given by the simple ratio of the number of firms that began the year with the same rating (B) and ended with an A rating to the total number of firms that began with a B rating.

More generally, we can let n_{ij} denote the number of individuals who were in state i in period $t-1$ and are in state j in period t . We can estimate the probability of an individual being in state j in period t given that they were in state i in period $t-1$, denoted by p_{ij} , using the following formula:

$$p_{ij} = \frac{n_{ij}}{\sum_j n_{ij}} . \quad (2)$$

Thus, the probability of transition from any given state i is equal to the proportion of individuals that started in state i and ended in state j as a proportion of all individuals in that started in state i .

Using the methods described above, it is possible to estimate a transition matrix using count data. Anderson and Goodman (1957) show that the estimator given in equation (2) is a maximum-likelihood estimator that is consistent but biased, with the bias tending toward zero as the sample size increases. Thus, it is possible to estimate a consistent transition matrix with a large enough sample. Moody's and Standard and Poor's, for example, provide

⁶ Including no change in credit rating.

estimates of transition matrices for different bond issuers, using observations on the individual transitions of thousands of different entities issuing bonds.⁷

B. Transition Matrices When Individual Transitions Unknown

As mentioned previously, the estimation of transition matrices is relatively simple when individual transitions are observed over time. Unfortunately, it is often the case that credit-quality transitions are imperfectly observed, and the best information available is an aggregate ratio or proportion showing the percent of total observations in a particular ratings category at a point in time. It is not possible to obtain maximum-likelihood estimates using the count method shown in equation (2) using such a time series of aggregate proportions data. However, if the time series of observations is sufficiently long, it is possible to estimate a transition matrix from aggregate data using quadratic programming methods.

Suppose that instead of observing the actual count of transitions from the different credit qualities, we only observe the aggregate proportions, $y_j(t)$ and $y_i(t-1)$, which represent the proportion of observations with credit quality j and i respectively. We can write a stochastic relation relating the actual and estimated occurrence of $y_j(t)$:

$$y_j(t) = \sum_i y_i(t-1)p_{ij} + u_j(t). \quad (3)$$

Following Lee, Judge, and Zellner (1970), we can write this equation in matrix form as follows:

$$y = Xp + u, \quad (4)$$

where

$$y = [y_1 \quad y_2 \quad \dots \quad y_{R-1}]' \\ = [y_1(1), y_1(2), \dots, y_1(T) \quad y_2(1), y_2(2), \dots, y_2(T) \quad \dots \quad y_{R-1}(1), y_{R-1}(2), \dots, y_{R-1}(T)]', \quad (5)$$

$$X_j = \begin{bmatrix} y_1(0) & y_2(0) & \dots & y_R(0) \\ y_1(1) & y_2(1) & \dots & y_R(1) \\ \vdots & \vdots & \ddots & \vdots \\ y_1(T-1) & y_2(T-1) & \dots & y_R(T-1) \end{bmatrix} \text{ for } j = 1, 2, \dots, R-1, \quad (6)$$

so that

$$X = \begin{bmatrix} X_1 & 0 & \dots & 0 \\ 0 & X_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & X_{R-1} \end{bmatrix}, \quad (7)$$

and

⁷ See Keenan, Hamilton, and Berthault (2000); and Standard and Poor's (2004) for examples.

$$p = [p_1 \quad p_2 \quad \dots \quad p_{R-1}]' \tag{8}$$

$$= [p_{11}, p_{21}, \dots, p_{R1} \quad p_{12}, p_{22}, \dots, p_{R2} \quad \dots \quad p_{1,R-1}, p_{2,R-1}, \dots, p_{R,R-1}]',$$

$$u = [u_1 \quad u_2 \quad \dots \quad u_{R-1}]' \tag{9}$$

$$= [u_1(1), u_1(2), \dots, u_1(T) \quad u_2(1), u_2(2), \dots, u_2(T) \quad \dots \quad u_{R-1}(1), u_{R-1}(2), \dots, u_{R-1}(T)]'.$$

Lee, Judge, and Zellner (1970, Chapters 1 and 3) suggest minimizing the sum of squared errors in equation (9) using the OLS method, subject to linear constraints on the transition probabilities, p . They note that OLS is equivalent to solving a quadratic programming problem. MacRae (1977) notes that the variance of the error term u depends on the magnitude of y_{t-1} , so using OLS estimation techniques will yield consistent but not efficient estimates. She demonstrates how to correct for the heteroscedasticity in the error term and produce a more efficient estimator using an iterative generalized least squares technique for calculating the matrix of transition probabilities, p . The first step in the procedure is to estimate the transition matrix, and then use this to calculate a consistent estimate of the conditional covariance matrix, denoted by Ω . The estimated covariance matrix is then used to obtain a subsequent estimate of the transition probabilities, and the procedure is repeated until convergence.⁸

Conditioning Transition Matrices on Fundamentals

Ideally, credit-transition matrices should be estimated using data from an entire economic or credit cycle, so that the estimated probabilities provide an accurate representation of average likelihoods, and are not unduly sensitive to the selection of the sample period. However, there may be circumstances where it is desirable to condition the transition matrices on particular macroeconomic variables or episodes. For example, if one is interested in the dynamics of credit quality during economic downturns, then the transition probabilities could be estimated using data that correspond to times when the economy was in recession. Bangia and others (2002) and Nickell, Perraudin, and Varotto (2000) find strong evidence of differences in transition matrices during periods of expansion and contractions for corporate bond issuers. They find downgrade probabilities are significantly higher but more stable during recessions, while upgrade probabilities are slightly lower. Their results are intuitively appealing, since it is inherently plausible that credit quality migrations would vary with the economic cycle, especially for lower quality credit risks.⁹ Credit-transition matrices can also be estimated for crisis periods or for periods before and after a major structural break, such as a major financial deregulation.

⁸ See Appendix I for further details.

⁹ Carty (1997), using a sample of Baa, Ba, and B-rated companies from 1920 to 1996, finds evidence that higher real GDP growth tends to lower the risk of default from lower rated (B) companies.

III. APPLICATION TO U.S. DATA

A. Nonperforming Loan Data

Section II demonstrated the feasibility of estimating transition matrices using aggregate proportions data. With this methodology in mind, we can now turn to a concrete application: the problem of understanding credit quality dynamics when data on individual transitions are not available. This situation arises in many countries, where often the only data that are available on credit quality are supervisory data on asset classifications (e.g., nonperforming loans).

Asset classification data have several shortcomings, including the fact that there is an upward bias from supervisors showing an overly optimistic assessment of bank balance sheets. Most supervisory data are based on book-value accounting instead of market-value accounting, so classification changes tend to lag behind the real evolution of credit risk, especially in deteriorating economic conditions. Loan classification data are also subject to regulatory forbearance and supervisory pressure, which may reduce their usefulness as an indicator of underlying credit quality. Despite these weaknesses, supervisory data are often the best available information for an entire financial system, since they are gathered and calculated by a single entity, usually using a consistent methodology.

One of the first difficulties in using supervisory data is that the data on nonperforming assets often do not include items that have been written off, or assets of institutions that have failed. Such datasets suffer from a survivor bias, because the estimated proportions of performing assets are biased upward by the exclusion of failed assets. Items in the “loss” category are thus systematically excluded, biasing the estimated parameters in the transition matrix. To obtain accurate estimates of a transition matrix, it is important that all failed assets and institutions are included in the sample. If it is not possible to observe failed institutions, it may be possible to estimate a transition matrix for survivors by only including observations on institutions that survive until the final period. Such a transition matrix would only be applicable to “survivors,” but it could be estimated on a consistent basis.

Another difficulty that can arise in applying the methodology to a particular dataset is the problem of sample selection. If the underlying structure of the financial system is undergoing profound structural changes (e.g., during the transition from a centrally planned economy to a market economy, or in the wake of widespread deregulation), then using a time-invariant transition probability model may not be appropriate.¹⁰ A short time horizon for the data sample is another problem that can arise when estimating transition matrices. Many supervisors have only recently begun gathering reliable data on nonperforming assets, and so it is often difficult to get long spans of data that are derived consistently. However, the

¹⁰ That is, a stationary Markov process. It is possible to estimate time-varying Markov transition matrices, but the additional data required to implement the technique are substantial. Since the object of this paper is to outline a methodology for use in situations with only limited data, the time-varying case is not considered here.

problem of short time horizons is less acute than the issue of structural change, since it affects the interpretation of results and not the validity of the underlying model.

FDIC Data

With the above caveats in mind, we turn now to an empirical application of the technique. We obtained quarterly data on nonperforming loans for the United States for the period 1984 to 2004. The data are taken from the FDIC's *Statistics on Banking*, and are derived from call reports to the Federal Financial Institutions Examination Council (FFIEC) and the Office of Thrift Supervision. The data cover all U.S. commercial banks that are insured by the FDIC.¹¹

Four proportions are calculated, expressed as a percentage of total loans, leases, and cumulative loan charge-offs:

- performing loans and leases;
- loans and leases past due 30–89 days;
- loans and leases past due 90 or more days, plus loans and leases in nonaccrual status; and
- cumulative charge-offs on loans and leases.¹²

Figure 1 presents a plot of the different categories of credit quality. We can see from this figure that there appears to be a structural break near the middle of the series, after which the ratios all become much more stable. Statistical tests¹³ reveal evidence of a structural break around the beginning of 1993. The impact of structural breaks is discussed later in this section.

Unconditional Estimates

Using the raw data depicted above, the GLS method of estimation described above was applied to the observations. Using the full sample, we are able to derive the quarterly transition matrix shown in (Table 1).¹⁴

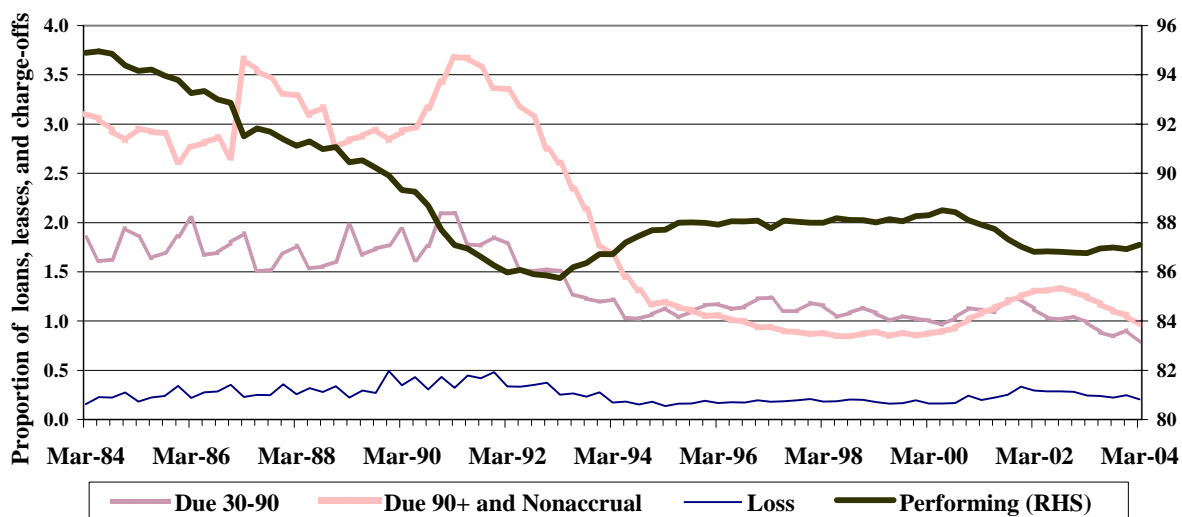
¹¹ See Appendix I for further details of the FDIC data.

¹² The loss category (loans and leases charge-offs) is defined an absorbing state, so all losses from the first time period should be included in this category, hence the cumulative sum of loan charge-offs is added to total loans and leases to derive the denominator in all of the proportions.

¹³ The tests outlined in Perron (1994) for a unit root with a structural break were applied to the series. The break point was identified by selecting the sample that corresponded to the regression with the highest t-statistic. The most common break point was identified as March 1993, followed by June 1993.

¹⁴ In all of the tables, the categories A, B, C, and D correspond to the categories of performing, due 30–89 days, nonaccrual plus due 90 days or more, and loss, respectively.

Figure 1. U.S. Nonperforming Loan Ratios



Source: FDIC, *Statistics on Banking*, Loans and Leases for All U.S. Commercial Banks Insured by the FDIC.

Table 1. Estimated Quarterly Transition Matrix for U.S. Commercial Bank Loans and Leases

(All Commercial Banks, 1984:1–2004:1)

		To			
		A	B	C	D
From	A	0.997	0.002	0.000	0.001
	B	0.000	0.852	0.067	0.080
	C	0.000	0.032	0.955	0.013
	D	0.000	0.000	0.000	1.000

Source: Author’s calculations.

The estimated transition matrix can be tested formally to ascertain whether the Markov structure is time invariant (i.e., whether the transition probabilities, p_{ij} , are constant over time). Kelton and Kelton (1984) propose a test for stationarity of the transition probabilities, based on estimating transition matrices for two subsamples and comparing them using an F test. When we apply this methodology to the estimated transition matrix, we find that we can reject the assumption that the transition probabilities are stationary. An examination of Figure 1 confirms this result, since it is clear that the second half of the sample appears to behave differently from the first half. Applying the methods discussed in Perron (1994), we find that the most common break date is around March 1993. When we re-estimate the transition probabilities using these two samples, shown in Table 2, we find that there is

indeed a significant difference in the transition matrices between the two periods. The proportion of performing loans improves significantly in the second period, most notably for loans in the 30–89 days overdue category, but also (weakly) for the loans in nonaccrual status and due more than 90 days.

Table 2. Estimated Quarterly Transition Matrix, Split Sample

1984:1–1993:1					1993:2–2004:1						
		T+1						T+1			
		1	2	3	4			1	2	3	4
T	1	0.994	0.006	0.000	0.000	T	1	0.997	0.002	0.000	0.001
	2	0.000	0.560	0.436	0.004		2	0.035	0.815	0.107	0.043
	3	0.062	0.083	0.755	0.101		3	0.000	0.000	0.886	0.114
	4	0.000	0.000	0.000	1.000		4	0.000	0.000	0.000	1.000

Source: Author’s calculations.

The change in the behavior of nonperforming loans between the two subsamples may also be due to changes in the underlying economic conditions. Indeed, when we examine the behavior of real GDP growth, shown in the dotted line in Figure 2, we can see that the deterioration in real GDP growth in the first half of the sample is matched by the decline in the performing loans ratio, with a lag of about 11 quarters.¹⁵ So, instead of breaking the sample in half, we can condition the transition matrices on the state of the economic cycle by constructing separate transition matrices for periods of low growth (defined as growth below the average trend rate for the period) and high growth (defined as growth above the average trend rate). The results of this approach are presented in Table 3.

We can see from these tables that there are significant differences in the transition matrices between periods of above average growth and below average growth. This result is consistent with the findings of Bangia and others (2002), who find significant differences between transition matrices over the business cycle.

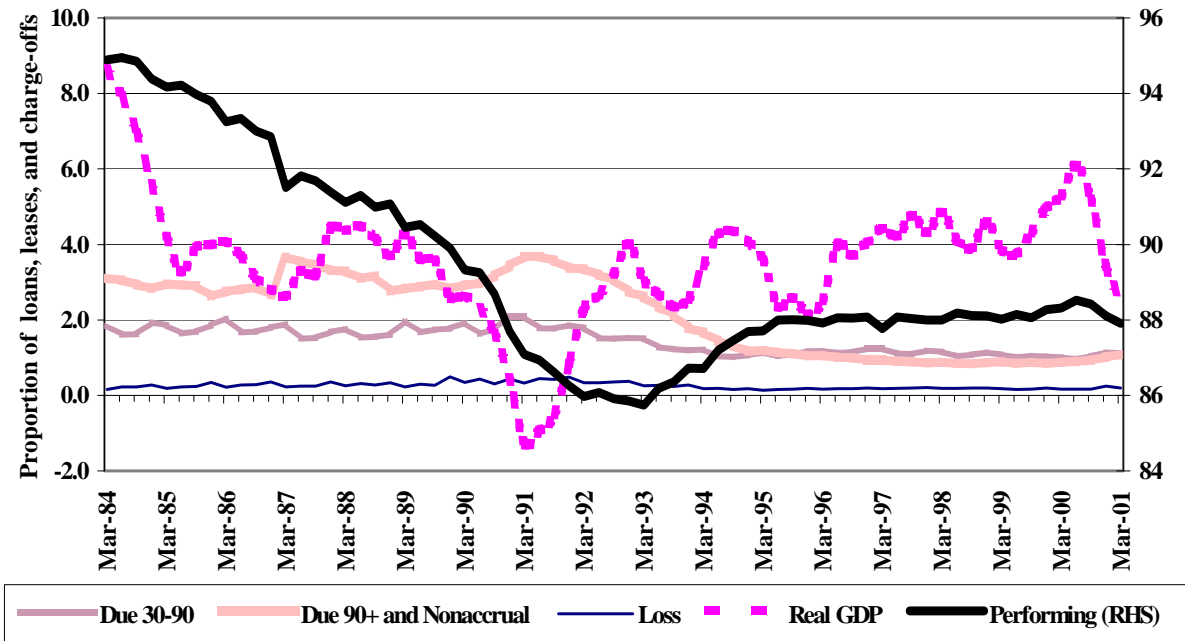
¹⁵ The correlation between changes in the percentage of loans that are performing and the year-on-year real GDP growth rate is 0.30, -0.19, -0.23, -0.31, -0.38, -0.38, -0.41, -0.39 for lags of 0, 6, 7, 8, 9, 10, 11, and 12 quarters, respectively.

Table 3. Estimated Quarterly Transition Matrix,
Low- and High-Growth Estimates

Above Average Growth					Below Average Growth						
T		T+1				T		T+1			
		1	2	3	4			1	2	3	4
T	1	0.995	0.006	0.000	0.000	1	0.992	0.002	0.000	0.006	
	2	0.000	0.651	0.080	0.270	2	0.000	0.799	0.201	0.000	
	3	0.000	0.000	0.945	0.055	3	0.070	0.056	0.874	0.000	
	4	0.000	0.000	0.000	1.000	4	0.000	0.000	0.000	1.000	

Source: Author's calculations.

Figure 2. Real GDP Growth and Loan Ratios



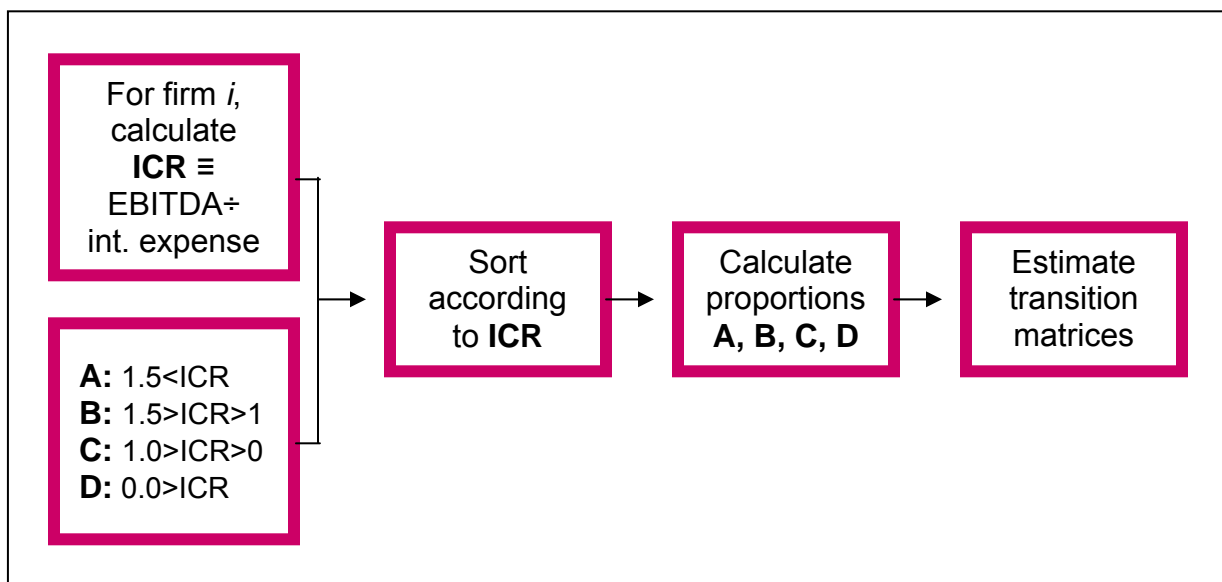
Sources: FDIC, *Statistics on Banking*, Loans and Leases for All U.S. Commercial Banks Insured by the FDIC; and U.S. Bureau of Economic Analysis, *National Income and Product Accounts Table*, Real Gross Domestic Product, year on year percentage change.

B. Interest Coverage Data

An alternative approach to estimating credit quality transition matrices is to use corporate sector balance sheet data on the ability of firms to meet their debt obligations. Several authors have used the interest coverage ratio (ICR, defined as the ratio of earnings before

interest, taxes, depreciation, and amortization to interest expenses) as a proxy for the underlying creditworthiness of a firm.¹⁶ For each firm we can estimate the ICR using information derived from their income statement. Depending on the value of the ICR, we can categorize each firm into one of four distinct groups. Then, by summing the interest expenses of firms in each category for a given point in time, we can get the proportion of total interest expenses in each category. Using these proportions, we can use the methodology described in Section II to estimate transition probabilities for a matrix like the one presented in Table 2.

Figure 3. Using Interest Coverage Data to Estimate Transition Matrices



Worldscope Data on the U.S. Corporate Sector

The data used to implement the approach outlined above are taken from the Worldscope database, and includes annual data for 1983–2003 on approximately 2,807 U.S. companies. The interest coverage ratio (ICR) is defined as the ratio of earnings before interest taxes depreciation and amortization, divided by interest expense on debt.¹⁷

Using the approach outlined in Figure 3, we arrive at the different proportions for interest expenses shown in Figure 4. These proportions can be used to estimate a transition matrix for the different ICR ratios. The results of this procedure are presented in Table 4. This table

¹⁶ See Goldman Sachs (1988, 2000); Andrade and Kaplan (1998); and Altman and Narayanan (1997).

¹⁷ See Appendix I for further details of the Worldscope data.

shows that the probability of upgrades (items below the diagonal) in the ICR of firms is higher than downgrades for the sample considered.

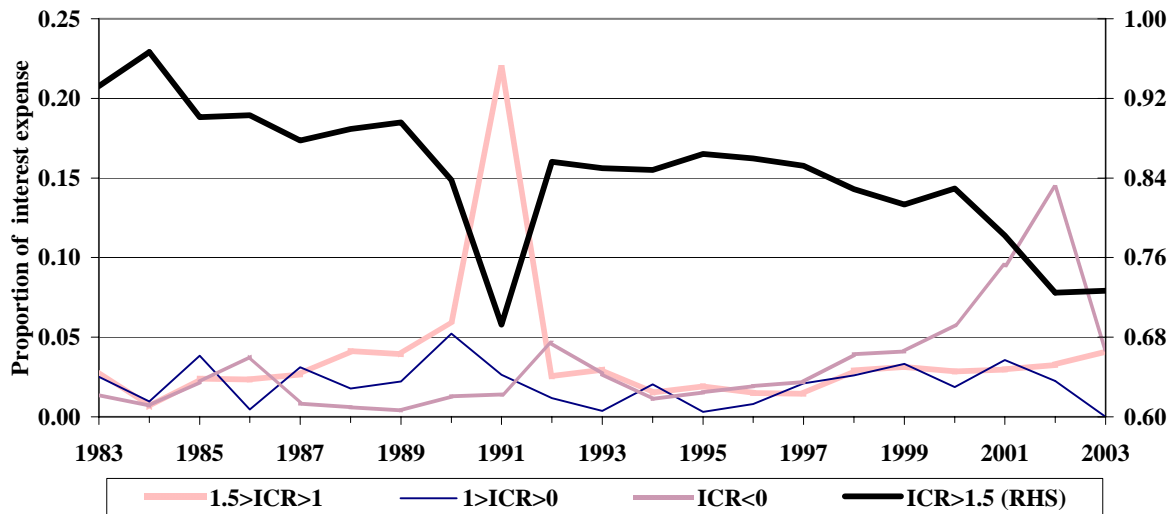
Table 4. Estimated Annual Transition Matrix Using Interest Coverage Ratio

		To			
		ICR>1.5	1.5>ICR>1	1>ICR>0	0>ICR
From	ICR>1.5	0.947	0.026	0.013	0.015
	1.5>ICR>1	0.357	0.547	0.041	0.055
	1>ICR>0	0.369	0.072	0.559	0.000
	0>ICR	0.000	0.000	0.000	1.000

Source: Author's calculations.

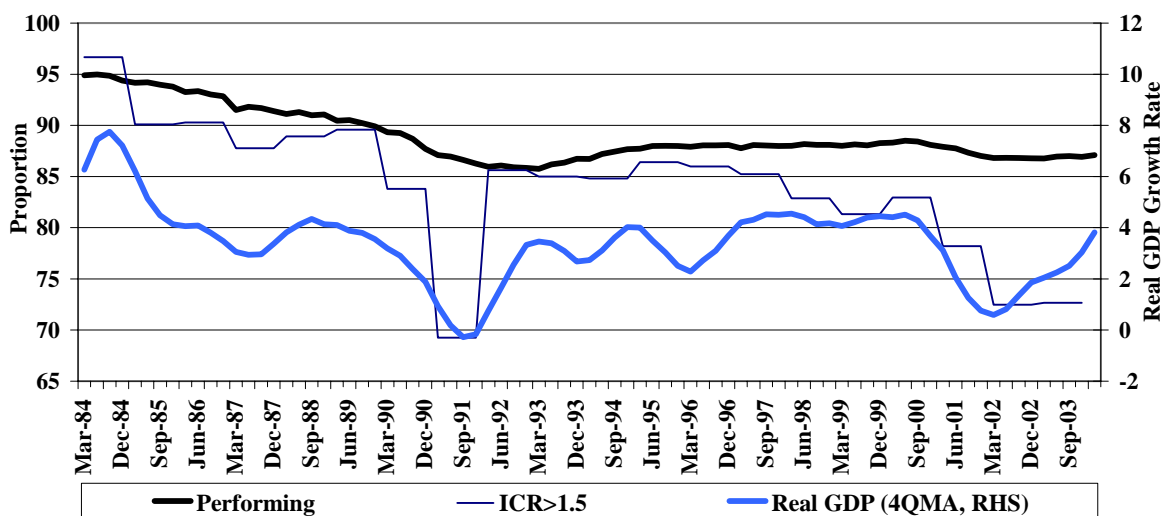
Figure 5 illustrates the close linkages between the interest coverage ratio, real GDP, and nonperforming loans for the U.S. economy. As we can see from this figure, the proportion of U.S. firms with the highest interest coverage ratio is closely related to the growth rate of the real economy. The proportion of performing loans also follows the path of real GDP, but with a lag of six to eight quarters.

Figure 4. Interest Coverage Ratio for U.S. Companies



Source: WorldScope, various years. The ICR is defined as earnings before interest, taxes, and depreciation (EBITDA) divided by the interest expense on debt. The ratios are defined as the sum of interest expenses for firms with interest coverage ratios in the indicated ranges as a proportion of total interest expenses of all U.S. firms in the sample.

Figure 5. Interest Coverage Ratio, Performing Loans, and Real GDP



Sources: FDIC, *Statistics on Banking*. Performing loans are defined as the proportion of all loans and leases that are not classified as past due or in nonaccrual status, for all U.S. commercial banks insured by the FDIC. U.S. Bureau of Economic Analysis, *National Income and Product Accounts Table*, Real Gross Domestic Product, four quarter moving average of year-on-year percentage change. WorldScope, various years. The ICR is defined as earnings before interest, taxes, and depreciation (EBITDA) divided by the interest expense on debt. The ICR>1.5 ratio is defined as the interest expenses of firms with an interest coverage ratio above 1.5 as a proportion of total interest expenses of all U.S. firms in the sample.

IV. ADDITIONAL APPLICATIONS

Having estimated a transition matrix to describe the evolution of credit risk, we can consider other applications of the technique. Bangia and others (2002) suggest using Hamilton’s (1989) switching regression framework to simulate paths of credit.¹⁸ For example, the transition matrix associated with contractions can be combined with information about the existing credit quality of a portfolio or system to simulate the impact of a prolonged recession on credit quality. In this manner we can use conditional transition matrices to estimate the impact of any future path of economic activity on credit quality. This approach can be used as inputs into stress tests of credit portfolios.

¹⁸ Hamilton proposes using a Markov switching model to estimate a nonlinear stationary process. Hamilton models the underlying process as if it is subject to discrete shifts in regime. The result of this procedure is a transition probability matrix that shows the estimated probability of going from an expansion state to either an expansion state or a contraction state, as well as the probability of going from a contraction state to either an expansion state or a contraction state.

V. CONCLUSIONS

This paper has outlined a way to estimate transition matrices for use in credit risk modeling with a decades-old methodology that uses aggregate proportions data. This methodology is ideal for credit-risk applications where there is a paucity of data on changes in credit quality, especially at an aggregate level. Using a generalized least squares variant of the methodology, this paper has provided estimates of transition matrices for the United States using both nonperforming loan data and interest coverage data. Consistent with other studies, this paper found evidence of transition matrices that vary over the economic cycle. In general, the methodology can be employed to condition the matrices on economic fundamentals and provide separate transition matrices for expansions and contractions, for example. The transition matrices can also be used as an input into other credit-risk models that use transition matrices as a basic building block.

TECHNICAL APPENDIX I

A. Estimating Transition Probabilities with Observed Transitions

Estimating a transition matrix is a relatively straightforward process, if we can observe the sequence of states for each individual unit of observation, i.e., if the individual transitions are observed. For example, if we observe the credit ratings of a group of firms at the beginning of a year and then again at the end of the year, then we can estimate the probability of moving from one credit rating to another.¹⁹ The probability of a firm having a particular credit rating at the end of the year, (e.g., A) given their rating at the beginning of the year (e.g., B) is given by the simple ratio of the number of firms that began the year with the same rating (B) and ended with an A rating to the total number of firms that began with a B rating.

More generally, we can let n_{ij} denote the number of individuals who were in state i in period $t-1$ and are in state j in period t . We can estimate the probability of an individual being in state j in period t given that they were in state i in period $t-1$, denoted by p_{ij} , using the following formula:

$$p_{ij} = \frac{n_{ij}}{\sum_j n_{ij}} . \tag{A1}$$

Thus, the probability of transition from any given state i is equal to the proportion of individuals that started in state i and ended in state j as a proportion of all individuals in that started in state i .

As a simple numerical example, illustrated in Table A1, consider a ratings company that observes a group of firms and their credit ratings at the beginning and end of a year. Suppose there are 100 firms with an A rating at the beginning of the year, of which only 70 remain in the A rating category. Suppose there are 100 firms with a B rating at the beginning of the year, with 15 of those firms ending up with an A rating. Similarly for the C rating, suppose there are 75 firms with a C rating at the beginning of the year, with only 10 moving to the A rating at the end of the year. Finally, for the default or D rating, suppose there are 50 firms at the beginning of the year, with 5 of those firms transitioning to an A rating. At the end of the year there are 100 firms with an A rating. Using equation (2), we can estimate the various transition probabilities, p_{iA} , as shown in the second column of the table. Using this approach, we can estimate the columns of the transition matrix shown in the right section of Table A1 for each credit rating A – D.

¹⁹ Including no change in credit rating.

Table A1. Illustrative Example of Using Count Proportions to Estimate Transition Probabilities

Count Data			Transition Probability Matrix					
	Number of firms in period t A=100	Probability of transition to A	To Rating					
				A	B	C	D	
Number of firms in period $t-1$	A=100	70	$p_{A,A} = 70/100$	From Rating A	0.70			
	B=100	15	$p_{B,A} = 15/100$	B	0.15			
	C= 75	10	$p_{C,A} = 10/ 75$	C	0.13			
	D= 50	5	$p_{D,A} = 5/ 50$	D	0.10			

Source: Author’s calculations.

B. Markov Probability Model²⁰

The first-order Markov probability model used in this paper is assumed to have the following characteristics. First, there is a population of individuals that moves among a finite set of R different states in a sequence of trials $t = 0, 1, 2, \dots T$. For a sample of size n from the population there are n units of observation that change over time according to independent and identically distributed time homogeneous Markov chains with R states. The discrete random variable x_t ($t = 0, 1, 2, \dots T$) can be used to represent the state of an individual in the population, and has a finite number of possible outcomes, denoted by s_i ($i = 1, 2, \dots R$).

In the context used in this paper, the population consists of the universe of loans (or assets) in a bank or a banking system. Each individual can be thought of as a one-unit loan, which can be classified into any of the R different categories of credit quality. Thus if there are loans worth \$1,000 outstanding, we can consider this as a population of 1,000 separate loans worth \$1 each, all with an associated credit quality. Growth in the total value of loans outstanding can be handled according to the procedures discussed in Kalbfleisch and Lawless (1984). They suggest a way of handling immigrations by assuming that additions to the population are initially made into the highest category and subsequently follow the same process as other members of the population. In the context of this paper, this is equivalent to assuming that new loans (or assets) are all performing initially, and then subsequently follow the same first-order Markov transition process as all other loans (or assets).

The probability distribution of credit quality depends only on the credit quality from the previous period, such that

$$\Pr(x_t | x_{t-1}, x_{t-2}, \dots, x_0) = \Pr(x_t | x_{t-1}), \quad \forall t. \tag{A2}$$

Thus, the probability of an ordered sequence or time series of credit quality can be written as

²⁰ This section is based on Lee, Judge, and Zellner (1970, Chapters 1, 3); and Kelton and Kelton (1984).

$$\Pr(x_0, x_1, x_2, \dots, x_T) = \Pr(x_0) \prod_{t=1}^T \Pr(x_t | x_{t-1}). \quad (\text{A3})$$

If $x_t = s_j$ and $x_{t-1} = s_i$, then we can write

$$\Pr(x_t = s_j | x_{t-1} = s_i) = p_{ij}(t) = p_{ij} \quad \forall t. \quad (\text{A4})$$

This formulation assumes that the Markov process is stationary (i.e., time invariant). Proceeding under this assumption, we can arrange the transition probabilities, p_{ij} , into an $(R \times R)$ transition probability matrix $P = [p_{ij}]$, which has the following properties

$$0 \leq p_{ij} \leq 1 \quad \sum_{j=1}^R p_{ij} = 1 \quad \text{for } i = 1, 2, \dots, R. \quad (\text{A5})$$

The summation condition above implies that the row sums must equal one.

Suppose that instead of observing the actual count of transitions from the different credit qualities, we only observe the aggregate proportions, $y_j(t)$ and $y_i(t-1)$, which represent the proportion of observations with credit quality j and i , respectively. We can write a stochastic relation relating the actual and estimated occurrence of $y_j(t)$:

$$y_j(t) = \sum_i y_i(t-1) p_{ij} + u_j(t). \quad (\text{A6})$$

We can write this equation in matrix form as follows:

$$y = Xp + u, \quad (\text{A7})$$

where

$$y = [y_1 \quad y_2 \quad \dots \quad y_{R-1}]' \quad (\text{A8})$$

$$= [y_1(1), y_1(2), \dots, y_1(T) \quad y_2(1), y_2(2), \dots, y_2(T) \quad \dots \quad y_{R-1}(1), y_{R-1}(2), \dots, y_{R-1}(T)]',$$

$$X_j = \begin{bmatrix} y_1(0) & y_2(0) & \dots & y_R(0) \\ y_1(1) & y_2(1) & \dots & y_R(1) \\ \vdots & \vdots & \ddots & \vdots \\ y_1(T-1) & y_2(T-1) & \dots & y_R(T-1) \end{bmatrix} \quad \text{for } j = 1, 2, \dots, R-1, \quad (\text{A9})$$

so that

$$X = \begin{bmatrix} X_1 & 0 & \dots & 0 \\ 0 & X_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & X_{R-1} \end{bmatrix}, \quad (\text{A10})$$

and

$$p = [p_1 \quad p_2 \quad \dots \quad p_{R-1}]' \quad (\text{A11})$$

$$= [p_{11}, p_{21}, \dots, p_{R1} \quad p_{12}, p_{22}, \dots, p_{R2} \quad \dots \quad p_{1,R-1}, p_{2,R-1}, \dots, p_{R,R-1}]'$$

$$\begin{aligned}
 u &= [u_1 \quad u_2 \quad \dots \quad u_{R-1}]' \\
 &= [u_1(1), u_1(2), \dots, u_1(T) \quad u_2(1), u_2(2), \dots, u_2(T) \quad \dots \quad u_{R-1}(1), u_{R-1}(2), \dots, u_{R-1}(T)]'.
 \end{aligned}
 \tag{A12}$$

Lee, Judge, and Zellner (1970, Chapters 1, 3) suggest minimizing the sum of squared errors in equation (A6) using OLS, subject to linear constraints. They note that OLS is equivalent to solving the following quadratic programming problem:

$$\begin{aligned}
 \text{Minimize } &_p \quad u' u &= (y - X p)'(y - X p) \\
 \text{subject to } & \sum_{j=1}^{R-1} p_{ij} &\leq 1 \\
 \text{and } & \sum_{j=1}^{R-1} p_{Rj} &= 0 \\
 \text{with } & p_{ij} &\geq 0.
 \end{aligned}
 \tag{A13}$$

We can write this in matrix form as follows:

$$\begin{aligned}
 \text{Minimize } &_p \quad u' u &= (y - X p)'(y - X p) \\
 \text{subject to } & G p &\leq \eta \\
 & p &\geq 0, \\
 \text{where } & G_{R \times R(R-1)} &= [I_1 \ I_2 \ \dots \ I_{R-1}] \\
 & \eta_{R \times 1} &= [1 \ 1 \ \dots \ 1 \ 0]'.
 \end{aligned}
 \tag{A14}$$

In this formulation, the matrix of constraints given by G is an $R \times R(R-1)$ matrix, composed of $R-1$ identity matrices of dimension R , and η is an $R \times 1$ column vector of ones, with a zero in the last row, to ensure that the loss state is absorbing.²¹ This formulation solves for the $R \times R$ unknowns using the system of $R \times T$ equations. Provided there are $T \geq R$ observations, it is possible to find a solution for P .²² The last row of the transition matrix is solved using the following equation:

$$p_{iR} = 1 - \sum_{j=1}^{R-1} \hat{p}_{ij}.
 \tag{A15}$$

MacRae (1977) notes that the variance of the error term u depends on the magnitude of y_{t-1} , so using OLS estimation techniques will yield consistent but not efficient estimates. She demonstrates how to correct for the heteroscedasticity in the error term and produce a more efficient estimator using an iterative generalized least squares technique for calculating the

²¹ See footnote 5. The final row of the transition probability matrix consists of zeros everywhere except for a 1 in the final column. Since we are using Equation (A13) to find the final column, then all of the other entries in the final row should be zero, hence the restriction on p_{Rj} for $j < R$.

²² This solution may be a local maximum if the problem is not strictly convex.

matrix of transition probabilities, p . The first step in the procedure is to estimate the transition matrix, and then use this to calculate a consistent estimate of the conditional covariance matrix, denoted by Ω . The estimated covariance matrix is then used to obtain a subsequent estimate of the transition probabilities, and the procedure is repeated until convergence. Using the same terminology as above, we can write the problem of minimizing a sum of squares subject to linear constraints on the probabilities as:

$$\begin{aligned}
 \text{Minimize}_p S &= (Y - \tilde{X} p)' \hat{\Omega}^{-1} (Y - \tilde{X} p) \\
 \text{where } Y_t &= [y_{1t} \quad y_{2t} \quad \dots \quad y_{R-t}]' \\
 Y &= [Y'_1, Y'_2, \dots, Y'_T]' \\
 X_t &= [y_{1t} \quad y_{2t} \quad \dots \quad y_{Rt}]' \\
 \tilde{X}_t &= I_{R-1} \otimes X_t' \\
 \tilde{X} &= [\tilde{X}'_0, \tilde{X}'_1, \dots, \tilde{X}'_{T-1}]' \\
 p &= [p_{11}, p_{21}, \dots, p_{R1} \quad p_{12}, p_{22}, \dots, p_{R2} \quad \dots \quad p_{1,R-1}, p_{2,R-1}, \dots, p_{R,R-1}]' \\
 \hat{\Omega}_t &= \text{diag}(p' X_t) - p' \text{diag}(X_t) p \\
 \hat{\Omega} &= \text{diag}[\hat{\Omega}_0, \hat{\Omega}_1, \dots, \hat{\Omega}_{T-1}] \\
 \text{subject to } Gp &\leq \eta \\
 p &\geq 0 \\
 \text{where } G_{R \times R(R-1)} &= [I_1 \quad I_2 \quad \dots \quad I_{R-1}] \\
 \eta_{R \times 1} &= [1 \quad 1 \quad \dots \quad 1 \quad 0]'
 \end{aligned}
 \tag{A16}$$

Kalbfleisch and Lawless (1984)²³ demonstrate how to estimate transition probabilities in cases where the population of individuals changes over time. In their formulation, immigration is assumed to occur into the highest category, with new entrants following the same transition matrix as the rest of the population. To handle immigration and emigration, the X_t matrix is redefined to consist of the ratio of the total number of observations in each class in the present period, $N_{i,t}$ divided by the total number of individuals in the population at the end of the previous period, n_{t-1} . The Y_t matrix consists of the ratio of the total number of observations in each class in the previous period before immigration occurs, $M_{i,t}$, again divided by the total number of individuals in the population at the end of the previous period, n_{t-1} . Thus, we replace X_t and Y_t in the previous equation with

²³ Kalbfleisch and Lawless (1984) use alternate symbols in their equations, replacing Y_t with Z_t , X_t with W_t , \tilde{X} with B , and Ω with Q .

$$\begin{aligned} X_t &= n_{t-1}^{-1} [N_{1t} \quad N_{2t} \quad \dots \quad N_{Rt}]' \\ Y_t &= n_{t-1}^{-1} [M_{1t} \quad M_{2t} \quad \dots \quad M_{R-t}]' \end{aligned} \quad (\text{A17})$$

$$\text{where } n_{t-1} = \sum_{j=1}^R N_{j,t-1}.$$

Kelton (1981) demonstrates the consistency of the estimated transition probabilities found as a solution to the quadratic programming equation in (A12). Kelton and Kelton (1984) provide test statistics that can be used to test for stationarity of the transition probabilities. Standard errors for the parameter estimates can be calculated using standard GLS techniques:

$$\begin{aligned} \hat{\sigma}_p &= \sqrt{s^2 (\tilde{X}' \hat{\Omega}^{-1} \tilde{X})^{-1}} \\ \text{where } s^2 &= \frac{(Y - \tilde{X} p)' \hat{\Omega}^{-1} (Y - \tilde{X} p)}{T - (R-1)(R-1)} \end{aligned} \quad (\text{A18})$$

C. Data Descriptions

FDIC Data on Nonperforming Loans

The data on nonperforming loans used in this paper are taken from the FDIC's *Statistics on Banking*,²⁴ and are derived from call reports to the Federal Financial Institutions Examination Council (FFIEC) and the Office of Thrift Supervision. The data cover all U.S. commercial banks that are insured by the FDIC, which includes commercial banks insured by the FDIC through either the BIF or the SAIF. These institutions are regulated by one of the three federal commercial bank regulators (FDIC, Federal Reserve Board, or Office of the Comptroller of the Currency) and submit financial reports to the Federal Reserve (state member banks) or to the FDIC (state nonmember banks and national banks). All financial data represent the consolidation of domestic and foreign operations, including operations in "Other Areas" (represented by Guam, Puerto Rico, U.S. Virgin Islands, and all other U.S. Territories and possessions).

Data on Loans and Leases Past Due 30–89 Days represents all loans and leases that are 30–89 days past due. *Loans and Leases Past Due 90 Days or More* represents all loans and leases that are 90 days or more past due. *Nonaccrual Loans and Leases* represents all loans and leases that (a) are maintained on a cash basis because of deterioration in the financial position of the borrower, (b) payment in full of interest and principal is not expected, or (c) principal or interest has been in default for a period of 90 days or more unless the obligation is both well secured and in the process of collection.

²⁴ Available on the web at <http://www2.fdic.gov/SDI/SOB/>.

The numerator used in all of the series is the sum of total loans and leases for all commercial banks, plus cumulative charge offs.²⁵ Performing loans are defined as the sum of all loans and leases less those 30–89 days past due, 90+ days past due, and loans and leases in nonaccrual status. Loss loans (the absorbing state) are defined as the cumulative charge offs.

GDP Data

The real GDP data used in this paper are from the U.S. Bureau of Economic Analysis, National Income and Product Accounts Table, Table 1.1.6, Real Gross Domestic Product, Chained Dollars, Billions of chained (2000) dollars, seasonally adjusted at annual rates.²⁶ The annual changes are calculated as year-on-year percentage changes (i.e., March 2003 over March 2002).

Worldscope Data on the United States Corporate Sector

The data used in this study are extracted from the Worldscope database, using Thomson One Banker. They cover approximately 2,500 U.S. companies for the period 1983–2003. The series extracted are described below.

Earnings Before Interest, Taxes And Depreciation (EBITDA) represent the earnings of a company before interest expense, income taxes, and depreciation. It is calculated by taking the pretax income and adding back interest expense on debt and depreciation, depletion and amortization and subtracting interest capitalized (Source Code 18198).

Interest Expense on Debt represents the service charge for the use of capital before the reduction for interest capitalized. If interest expense is reported net of interest income, and interest income cannot be found the net figure is shown. It includes interest expense on short term debt, interest expense on long term debt and capitalized lease obligations, amortization expense associated with the issuance of debt, and similar charges (Source Code 01251).

The classification of companies according to their interest coverage ratio is described in Figure 3. Since the lowest rating category (D) is assumed to be an absorbing state, once a company has a negative interest coverage ratio, it is assumed to remain in that category.

²⁵ Cumulative charge-offs are calculated by cumulating charge-offs from March 1984 onward.

²⁶ Available on the web at <http://www.bea.gov/bea/dn/nipaweb/TableView.asp?SelectedTable=6&FirstYear=2003&LastYear=2004&Freq=Qtr>. Data last revised on December 22, 2004.

REFERENCES

- Altman, Edward I. and Paul Narayanan, 1997, "Business Failure Classification Models: An International Survey," in *International Accounting and Finance Handbook*, ed. By Choi, Frederick D. S. (New York: John Wiley and Sons, 2nd ed.).
- Altman, Edward I. and Anthony Saunders, 1997, "Credit Risk Measurement: Developments Over the Last 20 Years," *Journal of Banking and Finance*, Vol. 21 (December), pp. 1721–42.
- Andrade, Gregor, and Steven N. Kaplan, 1998, "How Costly is Financial (Not Economic) Distress? Evidence From Highly Leveraged Transactions That Became Distressed," *Journal of Finance*, Vol. 53, no. 5 (October), pp. 1443–93.
- Bangia, Anil, Francis X. Diebold, Til Schuermann, A. Kronimus, and C. Schagen, 2002, "Ratings Migration and the Business Cycle, With Applications to Credit Portfolio Stress Testing," *Journal of Banking & Finance*, Vol. 26, no. 2/3, pp. 235–64.
- Caprio, Gerard, and Daniela Klingebiel, 1996, "Bank Insolvency: Bad Luck, Bad Policy, or Bad Banking," *Annual World Bank Conference on Development Economics*, 1996 (Washington: World Bank). Available via the Internet: http://www.worldbank.org/research/interest/prr_stuff/working_papers/abcde96.pdf.
- Goldman Sachs, 1998, *Asia Bank NPLs: How High, How Structural? Tying Estimates to the Real Sector*, Goldman Sachs Investment Research Report (Hong Kong SAR: Goldman Sachs).
- , 2000, *Asia Bank NPLs III: How High, How Structural?* Goldman Sachs Global Equity Research Report (Hong Kong SAR: Goldman Sachs).
- Hamilton, James D., 1989, "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle," *Econometrica*, Vol. 57, issue 2 (March), pp. 357–84.
- Kalbfleisch, J. D. and J. F. Lawless, 1984, "Least-Squares Estimation of Transition Probabilities From Aggregate Data," *Canadian Journal of Statistics*, Vol. 12, no. 3, pp. 169–82.
- , and W.M. Vollmer, 1983, "Estimation in Markov Models from Aggregate Data," *Biometrics*, Vol. 39 (December), pp. 907–19.
- Keenan, Sean C., David T. Hamilton, and Alexandra Berthault, 2000, *Historical Default Rates of Corporate Bond Issuers, 1920–1999*, Special Comment, January 2000 (New York: Moody's Investors Service). Available via the Internet: <http://www.moodyskmv.com/research/whitepaper/52453.pdf>.

- Kelton, W. David, and Christina M.L. Kelton 1982, "Advertising and Intraindustry Brand Shift in the U.S. Brewing Industry," *Journal of Industrial Economics*, Vol. 30, issue 3 (March), pp. 293–303.
- , 1984, "Hypothesis Tests for Markov Process Models Estimated from Aggregate Frequency Data," *Journal of the American Statistical Association*, Vol. 79, no. 388, pp. 922–28.
- Lee, T. C., G. G. Judge, and A. Zellner, 1970, *Estimating the Parameters of the Markov Probability Model From Aggregate Time Series Data* (Amsterdam: North Holland).
- Lindgren, Carl-Johan, Gillian Garcia, and Matthew I. Saal, 1996, *Bank Soundness and Macroeconomic Policy* (Washington: International Monetary Fund).
- MacRae, Elisabeth Chase, 1977, "Estimation of Time-Varying Markov Processes with Aggregate Data," *Econometrica*, Vol. 45, issue 1 (January), pp. 183–98.
- Nickell, Pamela, William Perraudin, and Simone Varotto, 2000, "Stability of Rating Transitions," *Journal of Banking & Finance*, Vo. 24, Issue 1–2 (January), pp. 203–27.
- Perron, Pierre, 1994, "Trend, Unit Root, and Structural Change in Macroeconomic Time Series," in *Cointegration for the Applied Economist*, ed. by Rao, B. Bhaskara, (New York: St. Martin's Press).
- Standard and Poor's, 2004, *S&P Quarterly Default Update & Rating Transitions*, Global Fixed Income Research, October (New York: Standard and Poor's). Available via the Internet: <http://www2.standardandpoors.com/spf/pdf/fixedincome/DefaultUpdate2004Q3.pdf>.