Journal of Real Estate Finance and Economics, 27:1, 5–23, 2003

A Proportional Hazards Model of Commercial Mortgage Default with Originator Bias

BRIAN A. CIOCHETTI
Department of Finance, University of North Carolina, Chapel Hill, NC, U.S.A.
E-mail: tony@unc.edu

YONGHENG DENG
School of Policy, Planning and Development, University of Southern California, Los Angeles, CA, U.S.A.
E-mail: ydeng@usc.edu

GAIL LEE
Credit Suisse First Boston, New York, NY, U.S.A.
E-mail: gail.lee@csfb.com

JAMES D. SHILLING
University of Wisconsin, Madison, WI, U.S.A.
E-mail: jshilling@bus.wisc.edu

RUI YAO
Department of Economics and Finance, Baruch College, New York, NY, U.S.A.
E-mail: Rui_yao@baruch.cuny.edu

Abstract

A proportional hazards model with competing risks is specified and is extended to correct for the possibility of originator bias. The model is used to examine the ability of option-theoretic models of mortgage pricing to forecast commercial mortgage defaults. Among the findings, those especially of interest include the influence of contemporaneous loan-to-value and debt-service-coverage ratios on commercial mortgage default probabilities. The paper also finds that option-theoretic models of mortgage pricing are quite capable of producing default estimates that fit the actual default rates well, especially when the model is corrected for originator bias.

Key Words: commercial mortgages, default, competing risks, hazard model, sampling bias

1. Introduction

In this paper we ask the following question, can option-theoretic models of mortgage pricing forecast commercial mortgage defaults? In order to test this hypothesis we set up an econometric model of commercial mortgage defaults that is predicated on option pricing theory. The model has several antecedents in the literature (see, for example, Ciochetti et al., 2002; Deng et al., 2002; Han and Hausman, 1990; and Suyoshi, 1992).[1] The model offers an opportunity to determine if option pricing theory can foretell commercial mortgage defaults over a period in which commercial mortgage default rates rose, reached a plateau, and then rose again for an unusual peak, thereafter falling back to all-time lows.

The present analysis differs from previous studies by explicitly correcting for originator bias. Traditionally, previous studies of mortgage default have focused on the determinants of default, attempting to measure the effects of loan characteristics, borrower information, property characteristics, and economic conditions on this transition probability. Very few of these studies have allowed for the possibility of originator bias in predicting default. Originator bias is a problem in all research utilizing mortgage loans originated by a single lender to determine default probabilities. Estimates of a prototypical loan's probability of default, computed from such time-series data, can in all likelihood overstate the probability of default if the lender in question has relatively loose underwriting standards, or it can understate the probability of default if the lender in question has relatively stringent underwriting standards. Correcting for this bias can thus be important.

Due to this problem, two models will be estimated: a model that includes only measures of loan characteristics, borrower information, property character-istics, and economic conditions on default, without correcting for originator bias; and a model that includes these variables and corrects for originator bias. The former model allows one to determine the relative effect of contemporaneous debt service coverage and loan-to-value ratios on commercial mortgage defaults without considering sampling bias. The latter model allows one to determine the relative effect of these variables and the potential biases related to the use of loan data originated by a single lender.

The remainder of the paper is organized into five sections. Section 2 deals with the exact specification of the model. The model is comprised of two fifth-order polynomial baseline hazard functions, one for default and the other for prepayment. The model is then estimated over the sample period 1974–1995, and the parameter estimates are used to produce *ex post* default forecasts through the end of 1999. The sample consists of 2043 loans, originated by a single large institutional multi-line insurance company lender. Because the sample is obviously not drawn randomly from the universe of all commercial mortgages originated, and because we wish to use the model to forecast industry-wide default rates, we are confronted with a potential sampling bias. Estimates of a prototypical loan's probability of default, computed from such time-series data, can in all likelihood lead to biased coefficient estimates. To eliminate this bias, we weight each observation in our sample based on how representative that observation type (category) is relative to its population equivalent. We then use a maximum likelihood estimation procedure to describe how quarter-to-quarter variation in debt service coverage and loan-to-value ratios affect the probability of default on commercial mortgages.

Section 3 describes the default data. We focus on commercial mortgage defaults over the time period 1974–1995 for two reasons: first, to study the responsiveness of commercial mortgage defaults to certain factors like loan-to-value ratios and debt-service-coverage ratios, one needs to look for well-defined movements of rise and fall. That is, defaults must be observed during revivals, expansions, recessions, and contractions in general commercial property markets. Basing current default rate estimates only on a strong commercial property market could lead to a significant lack of precision. Second, a key criteria that was used in selecting data for the analysis was that the information had to be available continuously over time. This explains why we focus on 1974 as the starting point for the analysis. This is the year in which our default rate data became available on a continuous basis. Also, it is preferable to have a database in which individual outstanding, prepaid and defaulted commercial mortgages are matched with loan-to-value and debt-service-coverage data over time. Previous researchers have generally been forced to relate the default incidence on individual commercial mortgages to variables observed at the time of origination, to the age of the mortgage, and to macro-economic data, or they have been forced to relate highly aggregate measures of default incidence to national macro-economic variables.

Sections 4 and 5 present the results and our forecasts. The results show that contemporaneous loan-to-value ratios are of less importance than contemporaneous debt service-coverage ratios in explaining commercial mortgage defaults, though they are of considerable significance in explaining quarter-to-quarter variations in default rates. Additionally, improved default predictions result by using an estimation technique that places greater individual weights on prototypical loans and lesser weights on uncharacteristic loans. Taken as a whole, the results reported here are quite consistent with the hypothesized determinants of the default and prepayment decision of commercial mortgage borrowers. The results further show that option-theoretic models of commercial mortgage pricing are in fact quite capable of producing default estimates that fit the observed default rates well. The forecasting errors emanating from our model are generally quite small, and the predicted estimates are quite close for all observations, except for the 1986–1990 time period. Conclusions are drawn in Section 6.

2. The modeling framework

A number of previous studies have used proportional hazards models to investigate the determinants of commercial mortgage defaults. These studies include Vandell et al. (1993), Follain and Ondrich (1997), and Pavlov (1999). In the current study, we focus on a proportional hazards model with competing risks. The model expresses the quarterly projected default and prepayment rate vectors as a function of a set of predetermined variables.[2] From the default rate vector, the model generates default rates for various commercial mortgage loan cohorts. These default rates are then aggregated to provide summary measures of default behavior for the commercial mortgage market on the whole.

Following are the right-hand side variables that were selected for the present analysis:

(1) *Debt service coverage ratio*. Debt service coverage ratio (DCR) refers to the ratio of net operating income (NOI) on a property to the annual debt service obligation on that same property. We include as right-hand side variables both DCR at loan origination, as well as contemporaneous DCR during the study period under examination. We obtain an approximation of the contemporaneous DCR by the following procedure. We use the increases or decreases in the National Council of Real Estate Investment Fiduciaries

(NCREIF) property appreciation series (stratified by eight geographic regions and four property types) to approximate value over time. Projected NOI is then defined as the estimated property value times the NCREIF income yield. This estimates what the property would have earned if it experienced the same income effects as properties in the NCREIF database. The contemporaneous DCR is the ratio of projected NOI to the annual debt service obligation.

(2) *Loan-to-value ratio*. Loan-to-value (LTV) is estimated as a contemporaneous variable. Calculating contemporaneous LTV begins with an estimate of the market value of the loan from the borrower's perspective. This is equal to the present value of the remaining debt service payments on the loan discounted at the going market interest rate. We follow the procedure described in Ciochetti (1993) to estimate the going market interest rate for each loan over time. The contemporaneous LTV is then the ratio of the projected market value of the loan to the estimated valuation of that property.

(3) *Interest rate spread*. This spread variable (CAL RATIO) is the percentage difference between the market and face values of the mortgage (market value/face value -1).

(4) *Loan type*. There are four loan type variables included in the model, loans with accrual provisions, loans with graduated/step payment covenants, those with amortization provisions, and those with changing coupon rates over time (the base case). From a credit perspective, fully amortizing loans are the least risky. As a loan amortizes, principal is paid down, thus reducing the indebtedness of the borrower and reducing the risk of default. Conversely, the riskiest loans are accrual loans and/or those with graduated payment provisions. Likewise, prepayment behavior may differ according to loan type.

(5) *Loan size*. Two dummy variables are used for loan size: medium and large, with small loan size being the base case.

(6) *Rollover dummy variable*. A 0–1 variable for loans that either prepay or default in close proximity (one quarter) prior to the scheduled balloon date of the mortgage. Typically, borrowers will try to refinance the outstanding balance into a new mortgage. If market conditions are stressful at this particular time, refinancing may be difficult and credit losses associated with default could be quite high.

(7) *Property type*. The present study examines the four major property classifications: Apartment, Industrial, Office, and Retail. Apartment and industrial properties are generally assumed to represent the least credit risk due to the fungible nature of the former, and the credit worthiness of tenants occupying the latter. Loans secured by office and retail properties are felt to offer higher credit risk than those secured by apartment and industrial properties. We include dummy variables for apartments, industrial properties, and office buildings. The base category is retail.

(8) *Borrower type*. We examine the impact that borrower characteristics may have on default probabilities by categorizing borrowers into four groups: individual borrowers, partnership entities, corporate borrowers, and others. Our priors would suggest that individual borrowers, who are generally smaller in size, and take an active role in property management, may have stronger economic ties to the asset, and thus represent a lower risk to lenders. Additionally, individual borrowers may have a more vested interested in the property because the asset represents a larger proportion of their wealth. Corporate status, on the other hand, may protect the borrower somewhat more from the consequences of default. This may reduce the transaction costs of default, implying that corporate borrowers may default whenever equity is negative. Individual borrowers, partnerships, and corporations are included in the estimation, with other borrowers as the base category.

These variables were selected in part as a result of our desire to specify an empirical model of default and prepayment as parsimonious in the parameters to be estimated as possible. Variables (1)–(3) are measured on a contemporaneous basis, and account for the fundamental causes of default. The DCR may change over time due to a decrease in rents or an increase in vacancy rates. The LTV, on the other hand, may change over time because of a change in investors' risk preferences. We also measure DCR at origination to test if there is default information imbeded in the underwriting process.[3] Variable (3) measures the potential borrower savings from refinancing by capturing the effects of both the spread between the mortgage coupon and the prevailing market rate and the interaction between the spread and the remaining term of the mortgage. Variables (4)–(6) are related to the mortgage debt and are essential for valuing the default option. Variable (7) is related to the asset securing the mortgage. This variable is our proxy for the promised cash flows of the underlying mortgage. Variable (8) is related to borrower status. Our priors are that borrower types should have an impact on default rates.

Since both a call and put option are normally embedded in commercial mortgage debt, each of the default variables listed above is linked to the value of the prepayment option, and each of the prepayment variables is linked to the incentive to default.[4] One expects, for example, defaults to occur when the value of the property falls below the unpaid balance of the mortgage. However, a decline in property prices that would otherwise trigger a default might not if interest rates have risen and mortgage values have fallen. In this case, most borrowers would wish to maintain what is now a low-rate mortgage (i.e., slow down the speed of repayment). Similarly, when interest rates drop and property value and NOI increase, the borrower has an incentive to prepay. Sometimes the borrower can default, go through foreclosure, pay off the face value of the debt with the proceeds of the refinancing, thereby avoiding a prepayment penalty (Ciochetti and Riddiough, 1999).

Maximum likelihood estimation is then used to obtain estimates of all the equations in the system. The difficulty with the system estimation is that the individual parameter estimates (by construction) are sensitive to the specification of the entire model system. A serious specification error in one equation can therefore affect the parameter in both equations of the model. Yet neglecting the likelihood function of one or the other competing risk in the model can lead to biased estimates (toward zero).

Another problem—but one that can be adjusted for—is the problem of originator bias. When we have a sample of commercial mortgage loans originated by a single lender, we may end up with data that are thinly represented for certain categories of commercial mortgages. Consequently, when we estimate a proportional hazard model using these data, we can expect that the variance of the residuals for these thinly represented observations may be different from the residuals for typical observations in the population. This situation represents a correctable heteroskedasticity condition.

To construct a likelihood function for the fully specified competing risk proportional hazard model with a single-lender originator bias, we make two assumptions. First, we assume that the probability density of default and prepayment for commercial mortgage i at time t is

$$H^d_{i,t}(x_{i,t};\beta_d) = \exp(\gamma^d(t) + x'_{i,t}\beta_d), \tag{1}$$

$$H^p_{i,t}(x_{i,t};\beta_p) = \exp(\gamma^p(t) + x'_{i,t}\beta_p), \tag{2}$$

where $x_{i,t}$ is a vector of covariates (or regressors), either time-constant or time-varying, β_k is a vector of coefficients, and $\gamma^k(t)$ is the (log of) integrated baseline hazard rate for risk type k between $t-1$ and t. We model $\gamma^k(t)$ as a fifth-order polynomial function. This allows us to avoid imposing any restrictive functional form assumptions on the baseline hazard.

Second, we assume that a particular commercial mortgage at time t is a member of category k $(k=1,\ldots,K)$, and category k commercial mortgages constitute $p_{k,t}$ percent of the sample. We further assume that category k commercial mortgages represent $p^*_{k,t}$ of all commercial mortgages at time t within the population. The condition $p_{k,t} > p^*_{k,t}$ would therefore suggest that category k commercial mortgages are over-represented in the sample. To adjust for this over-representation, we can weight all category k commercial mortgages by a factor of $p^*_{k,t}/p_{k,t}$. This will result in a weighted sample that more accurately represents the population. Therefore, the sample maximum likelihood estimates that are applied to the population should more effectively represent the aggregate default rate for the population.

The joint survival function for observed outcomes depends on the probability that a commercial mortgage with characteristics $x_{i,t}$ is observed transiting from good standing to foreclosure, or from good standing to prepaid. The joint distribution is

$$S(t_d, t_p|x, \theta_d, \theta_p) = \exp\left(-\theta_d \sum_{t=1}^{t_d} \exp(\gamma^d(t) + x'_{i,t}\beta_d) - \theta_p \sum_{t=1}^{t_p} \exp(\gamma^p(t) + x'_{i,t}\beta_p) \right), \tag{3}$$

where t_d and t_p represent the duration of the mortgage until it is terminated by default or prepayment, respectively, and θ_d and θ_p are location parameters.

We assume that t_d and t_p are related by way of their unobserved determinants being related. The observed (uncensored) duration is $t = \min(t_d, t_p)$. We then write the probability of mortgage termination as follows. First, we write the probability of mortgage termination by default in period k as

$$\begin{aligned} F_d(k|\theta_d, \theta_p) = {} & S(k, k|\theta_d, \theta_p) - S(k+1, k|\theta_d, \theta_p) \\ & - 0.5\{S(k, k|\theta_d, \theta_p) + S(k+1, k+1|\theta_d, \theta_p) \\ & - S(k, k+1|\theta_d, \theta_p) - S(k+1, k|\theta_d, \theta_p)\}. \end{aligned} \tag{4}$$

The probability of mortgage termination by prepayment is

$$\begin{aligned} F_p(k|\theta_d, \theta_p) = {} & S(k, k|\theta_d, \theta_p) - S(k+1, k|\theta_d, \theta_p) \\ & - 0.5\{S(k, k|\theta_d, \theta_p) + S(k+1, k+1|\theta_d, \theta_p) \\ & - S(k, k+1|\theta_d, \theta_p) - S(k, k+1|\theta_d, \theta_p)\}. \end{aligned} \tag{5}$$

The probability that a mortgage will survive till period k and be right-censored after that is

$$F_c(k|\theta_d, \theta_p) = S(k, k|\theta_d, \theta_p). \tag{6}$$

The likelihood function is now specified as

$$\log L = \sum_{i=1}^{N} \{I_{di} \log(F_d(k_i)) + I_{pi} \log(F_p(k_i)) + I_{ci} \log(F_c(k_i))\}, \tag{7}$$

where N is the total number of observations, and I_{di}, I_{pi} and I_{ci} are indicator functions that take on the value of one if the ith loan is terminated by default, prepayment, or censoring, respectively, and zero otherwise. The core assumptions are that all borrower heterogeneity is captured by the set of covariates included in the analysis, and relative risks remain constant over the duration of the mortgage.[5]

Let us now go back to the discussion regarding originator bias. If we are concerned with estimating unknown parameters β_d and β_p for a sample of N loans originated by a single-lender, we could easily choose β_d and β_p jointly in order to maximize (7). However, if the sample estimates are to be applied directly to the population, and if there are significant differences in the levels of categorical similarities between the sample and the population, then we have a form of sampling design bias that can create a loss in estimation efficiency. Furthermore, applying these sample parameter estimates directly to the population can lead to systematic prediction errors.

In order to obtain more reliable and accurate forecasts, we proceed as follows. First, as outlined above, we weight each observation based on how representative that observation type (category) is relative to its population equivalent. Then we choose β_d and β_p jointly to maximize the following the likelihood function:

$$\log L = \sum_{i=1}^{N} w_i \{ I_{di} \log(F_d(k_i)) + I_{pi} \log(F_p(k_i)) + I_{ci} \log(F_c(k_i)) \}, \tag{8}$$

where w_i is an observation weight based on similarity criteria. By using population parameters as the basis for the relative weighting, we are shifting the major prepayment and default influencing characteristics in the sample to a point where they more closely approximate the population's. Hence, from a prediction standpoint, the model should have overall predictive abilities equal to, or superior to, the model given in (7). Also, because extreme and unrepresentative observations now have only a small influence on the likelihood function given in (8), the estimates of β_d and β_p should be more robust and efficient than those obtained by estimating (7).

3. Data and sample characteristics

The model is estimated using data on loans originated by a large institutional multi-line insurance company lender. All loans employed in the study are permanent loans secured by commercial real estate and originated over the period 1974–1990. In sum, 2,043 loan are used in the study. The data cover four major property types and nine geographic regions. Each loan is tracked from its origination date through prepayment, default, or the censoring point of the study as of the end of 1995.

In this study, we define default as the first event representing the foreclosure of the borrower's interest in the property; either onset of the foreclosure process, or in the case where foreclosure occurs rapidly, completion of the process. Of the 422 loans that are categorized as defaults, 360 are tracked as of the onset of the foreclosure process, while 62 complete the foreclosure process within the quarterly tracking period.[6]

Table 1 shows the number of loans originated in each of the years sampled between 1974 and 1990. The smallest loan cohorts are 33 and 34 loans, originated in 1989 and 1990, respectively. This coincides with the general downturn of the real estate markets in the late 1980s. The largest loan cohort consists of 231 loans originated in 1978.

The sample is geographically diverse within the United States. Table 2 shows the distribution of loans by geographic region. The highest volume of loans are originated in the Pacific region, with 516, followed by the South Atlantic region with 433 loans, and the East North Central region with 391 loans. Combined, these three regions represent approximately 65 percent of the sample. The East South Central region of the country experienced the lowest loan activity over the study period, with 20 loans.

Also shown in Table 2 is the distribution of loans by property type securing the loan. Notice that the sample is dominated by loans secured by office properties, with 888, or nearly 43 percent of the total. Loans secured by apartment and industrial properties are nearly equal, at 414, and 425 loans, respectively. The average loan size in the sample is

Table 1. Loan origination activity by year of origination.

Year of Origination	Number of Loans	Percent
1974	59	2.9
1975	59	2.9
1976	85	4.2
1977	198	9.7
1978	231	11.3
1979	198	9.7
1980	150	7.3
1981	84	4.1
1982	52	2.5
1983	177	8.7
1984	114	5.6
1985	190	9.3
1986	146	7.1
1987	140	6.9
1988	93	4.6
1989	33	1.6
1990	34	1.7
Total	2,043	100.0

Table 2. Sample characteristics.

	Number of Loans	Percent
Region		
East North Central	391	19.1
East South Central	20	0.9
Mid Atlantic	137	6.7
Mountain	216	10.6
Northeast	85	4.2
Pacific	516	25.3
South Atlantic	433	21.2
West North Central	99	4.9
West South Central	146	7.2
Total	2,043	100.0
Property type		
Apartment	414	20.3
Industrial	425	20.8
Office	888	43.5
Retail	316	15.5
Total	2,043	100.0

Variable	Mean	Std. Dev.	Min.	Max.
Loan size	$6,742,964	$15,366,412	$67,545	$370,000,000
LTV	72.1%	0.058%	5.6%	97.4%
DSCR	1.26	0.16	0.54	4.12

approximately \$6.7 million with a range of \$65,000 to \$370 million. The loans originated in the sample bear an initial LTV of 72 percent and an average initial DCR of 1.26.

Since we lack ongoing income and property values at the individual loan level, we estimate contemporaneous LTVs and DCRs by using quarterly data as secured from both the NCREIF, as well as the American Council of Life Insurance (ACLI). NCREIF reports not only total return indices, but also property value and income (NOI) indices. We use the NCREIF property value indices, as stratified by property type and region in order to estimate contemporaneous property values, which approximate the price paths of individual loans in the sample.

In order to estimate contemporaneous loan values, we first employ ACLI mortgage commitment data to fit a third-order polynomial function to contemporaneous mortgage contract rates, using remaining loan term to account for the term structure effect associated with each individual mortgage. We also fit a regression model in order to estimate a mortgage spread to account for variation by property type and region of loan origin. This modeled spread is then added to the fitted mortgage rate to create a contemporaneous mortgage contract rate. This rate is then used to discount remaining contractual loan payments in order to derive the contemporaneous market value of each loan.

To estimate contemporaneous DCRs, we match individual property values as described above with the quarterly income return series as provided by NCREIF to derive an estimate of net operating income. As with the property indices, the income series is stratified by both region of loan origin as well as property type securing the loan. Annual debt service is derived from loan terms describing the contractual cash flows due on each underlying mortgage. The ratio of NOI to scheduled annual debt service payments provides an estimate of the contemporaneous DCR.

4. Empirical estimates

4.1. Unweighted results

The results of maximizing (7) are provided in Table 3. Column (1) are the estimates of the prepayment model. Column (2) are the estimates of the default model.

A few brief comments about Table 3 seem in order. First, we should observe an inverse relation between commercial mortgage default and contemporaneous DCR—the lower the contemporaneous DCR, the higher the probability of default prior to maturity. The model confirms this: the coefficient on the contemporaneous DCR variable (DCR) is negative and significant at the 5 percent level. The DCR at origination, denoted as ORIGDSC, shows up positive in the default hazard function, but insignificant.

Second, we would expect to observe a positive relation between default rates and LTV conditional on debt-service-coverage ratio. That is, for a property owner to exercise his or her default option, two conditions must generally hold: (1) the market value of the property must be less than the market value of the debt, and (2) net operating income must be less than the current period's scheduled mortgage payment; otherwise, default should not occur. Also, the nature of this relationship ought to be nonlinear. In other words, defaults

Table 3. Maximum likelihood estimates for competing risks hazard model of prepayment and default without ACLI weighting.

	Model			
	Prepayment		Default	
Variables	Parameter estimate	*T*-ratios	Parameter estimate	*T*-ratios
Medium	– 0.2125	– 1.6909*	0.5110	3.2454***
Large	– 0.4590	– 2.4032**	0.9070	5.1829***
AMZDUMMY	– 0.0773	– 0.3171	– 0.9041	– 6.9570***
ACRDUMMY	– 1.1586	– 1.5548	0.5648	2.3268**
GPMDUMMY	– 0.0220	– 0.1185	0.2167	1.5989
Apartment	– 0.0206	– 0.1196	– 0.0537	– 0.2791
Industrial	0.0150	0.0964	0.0977	0.5488
Office	– 0.0377	– 0.2250	– 0.1386	– 0.6228
Individual	0.0907	0.3875	– 0.0572	– 0.2152
Partnership	0.2841	1.5243	0.0996	0.5835
Corporation	– 0.0213	– 0.0996	– 0.1629	– 0.7931
ORIGDSC	– 0.6538	– 1.4522	0.4584	1.1074
LTV ratio	1.6179	1.6712*	– 0.5589	– 1.0602
CAL ratio	15.8263	5.1648***	7.2649	3.3668***
DCR	0.5963	3.8285***	– 0.8026	– 3.7824***
Balloon	3.9973	21.4998***	1.7127	11.7790***
LTV squared	– 2.1845	– 3.1420***	0.5013	3.0223***
CAL squared	– 23.2005	– 1.7281*	– 42.6309	– 3.3772***
LOC	4.34E-05	0.8748	1.59E-04	0.9423
Log likelihood	– 4259.0971			

Note. The model is estimated by maximizing the likelihood function of the competing risk hazard model of prepayment and default. Two fifth-order polynomial functions are estimated as the baseline hazard functions (not reported). *, ** and *** indicate significance at 10%, 5%, and 1% levels, respectively.

should increase, at an increasing rate, with the loan-to-value ratio for loans with marginal equity and for loans with increasingly negative equity. The estimates in Table 3 do indeed suggest that there is a positive and nonlinear relationship between commercial mortgage defaults and contemporaneous LTV.

Third, as might be expected, the interest rate spread variable (CAL RATIO) is highly significant in explaining commercial mortgage default and prepayment. The model suggests that borrowers react proportionately more to positive spreads, caused by falling interest rates, than to negative ones—a nonlinear response. For example, if the mortgage coupon rate is 9 percent and the prevailing rate on new mortgages is 8 percent, the 1 percentage point positive spread provides a somewhat modest incentive to refinance. And as the spread between the mortgage coupon and the prevailing market rate widens, the incentive to refinance continues to grow. Beyond a certain point, however, prepayments decline.[7] In terms of default, a high interest rate spread moderates the risk of default. The positive spread lowers defaults on commercial mortgages, because existing loans now have below-market rates.

Fourth, balloon mortgages default more often than fully amortizing loans. There can be a variety of reasons why this occurs. The balloon date may represent a crisis point for the borrower. NOI might have decreased, for example, or interest rates may have increased, leading to an insufficient coverage ratio. Or perhaps property values may have decreased, thereby preventing refinancing. Whatever the reason, as the balloon date approaches, the option to default later on because the borrower has an underpriced option diminishes, making default the optimal decision to exercise.

Fifth, the set of 0–1 dummy variables for medium and large borrowers, amortization, denoted AMZDUMMY, accrual loans, denoted ACRDUMMY, and step-rate loans, denoted GPMDUMMY, are all significant at the 1 percent level, suggesting that borrower type and loan type are quite important in forecasting defaults. The results suggest, for example, that large and medium-sized borrowers default more than small borrowers, and that accrual and step-rate loans have higher default rates than interest-only loans (the base category).

The last point to note about these estimates is that property type does not seem to affect default risk; although, property type does seem to affect prepayments, particularly apartments and offices.

Table 4 presents simple default probabilities calculated separately for each category shown. The values are cumulative estimates of commercial mortgage defaults and indicate the proportion of commercial mortgages terminating by various mortgage durations. For

Table 4. Estimates of cumulative default probabilities by specific durations by selected mortgage and borrower characteristics (without ACLI weighting).

	Predicted by Model		
Characteristic	5 years	10 years	15 years
LTV category:			
65%	2.94	9.53	20.59
75%	3.05	9.88	21.29
85%	3.21	10.36	22.25
DCR category:			
1.15	2.84	10.44	24.21
1.25	2.51	9.51	22.21
1.50	2.01	7.51	17.81
Large	1.86	6.05	13.24
Apartment	2.87	9.22	19.76
Industrial	2.97	9.54	20.40
Office	2.82	9.07	19.46
Individual	3.21	10.25	21.82
Partnership	3.88	12.30	25.81
Corporation	2.87	9.22	19.75

Note. Differences between the default probabilities for the categories are statistically significant at the 0.10 level (or better) only for LTV, DCR, and large loans.

example, 10.44 percent of the commercial mortgages with debt service coverage ratios of 1.15 will default by year 10; this proportion increases to 24.21 percent by the end of year 15. For commercial mortgages with debt service coverage ratios of 1.50, the proportion that will default by the end of year 15 is 17.81. Similarly, for commercial mortgages with loan-to-value ratios of 85 percent, 10.36 percent of these mortgages will default by year 10 and 22.25 percent will default by year 15. This is the relationship predicted by standard option-theoretic models of commercial mortgage pricing.

4.2. Weighted results

Estimates of (8) are provided in Table 5. All observations have been weighted to adjust for sampling bias and all standard-error estimates have been modified for design effects. Column (1) are the estimates of the prepayment model. Column (2) are the estimates of the default model.

Table 5. Maximum likelihood estimates for competing risks hazard model of prepayment and default with ACLI Weighting.

	Model			
	Prepayment		Default	
Variables	Parameter Estimate	*T*-ratios	Parameter Estimate	*T*-ratios
Medium	– 0.3604	– 2.4031**	0.4545	2.7739***
Large	– 0.4485	– 2.2747**	0.9821	5.7007***
AMZDUMMY	– 0.1219	– 0.4466	– 0.6427	– 5.2554***
ACRDUMMY	– 2.3818	– 1.4720	0.9800	4.0914***
GPMDUMMY	– 0.0219	– 0.1056	– 0.2619	– 1.9473***
Apartment	0.1201	0.5659	0.1079	0.5357
Industrial	0.0627	0.3865	0.2910	1.7999
Office	0.1126	0.6382	0.3423	1.7118
Individual	0.1662	0.5691	0.2923	1.1290
Partnership	0.3655	1.6189*	0.1819	1.1038
Corporation	– 0.0656	– 0.2604	– 0.0248	– 0.1236
ORIGDSC	– 0.7853	– 1.8137*	0.7071	2.6518
LTV ratio	– 0.7662	– 0.6642	0.1918	0.3854
CAL ratio	14.7331	4.0746***	4.3649	2.2079***
DCR	0.3513	1.7685***	– 1.0577	– 5.0414***
Balloon	4.1676	18.7759***	1.6858	12.7740***
LTV squared	– 1.0799	– 1.3590**	0.3069	1.8781***
CAL squared	– 16.6128	– 1.0868*	– 28.2562	– 2.2712***
LOC	1.25E-04	0.8453	6.95E-05	1.0211
Log likelihood	– 4072.0606			

Note. The model is estimated by maximizing the likelihood function of the competing risk hazard model of prepayment and default. Two fifth-order polynomial functions are estimated as the baseline hazard functions (not reported). *, ** and *** indicate significance at 10%, 5%, and 1% levels respectively.

We obtain the population weight by matching loans in ACLI's survey of commercial mortgage lending activity with those in our sample. Our matching algorithm is as follows. We start by counting newly originated loans in both ACLI's survey and in our sample in a particular quarter, property type and region category. We then assign the counts from the larger ACLI pool equally to the loans in our sample matched by quarter, property type, and region as a raw weight. For quarters that the lender in our sample did not originate any new loans, we assign the weights equally to the loans in the quarter immediately before or after. If there are still no matching loans in our sample after applying the above criteria, we go as far as two quarters back and two quarters ahead to find a match. We are able to match 16,210 loans in the ACLI survey with 2,016 loans in our sample. The raw weights for individual loans in our sample range from 1 to 100, reflecting a potentially significant originator bias in certain circumstances. The ACLI survey has a total of 20,507 loans originated between 1974 and 1990.

With respect to the reported results in Table 5, the coefficient on the balloon dummy variable changes little from that in Table 3, but that on the interest rate spread, LTV, and DCR variables change dramatically. The coefficients on the interest rate spread and LTV variables in Table 5 are about two-thirds as large as those in Table 3, and that on the DCR variable is about 1.3 times as large.

Note also that the coefficient on loans with graduated/step payment covenants differs considerably from that in Table 3. Additionally, the estimates in Table 5 generally explain the default rates on industrial and office loans better than the estimates in Table 3. The estimates suggest that property type may matter in explaining default, but not in explaining prepayments.

Table 6 indicates the cumulative default probabilities for the weighted model. The probabilities are easily calculated from equation (1) by setting the values for the covariates at their means. From Table 6 it can be seen that when the relative risks are high, the cumulative default probabilities are about 8–18 percent by year 15. As can be seen, these default probabilities are much lower than those reported in Table 4. The differences underscore the distinct character of our sample of 2,043 loans and that of the loans in ACLI's survey of commercial mortgage lending activity. The loans in our sample on the whole are made up of more office and industrial loans (which tend to be riskier), and a larger number of loans originated in the East North Central, Pacific, and South Atlantic. In contrast, the loans in ACLI's survey of commercial mortgage lending activity are made up of more apartment and retail loans, and a larger number of loans originated in the Mid Atlantic, Mountain, and Central.

There is also strong evidence that the hazard increases much more significantly with a decrease in DCR than with an increase in LTV. We have grouped loans by DCR: 1.15, 1.25, and 1.50. We have also grouped loans by LTV: 65 percent, 75 percent, and 85 percent. We find that a decrease in DCR has a much larger impact on cumulative default probabilities than an increase in LTV (see Table 7). These results suggest that the ability to meet debt service obligation is the most relevant measure of commercial mortgage default. This may indicate that defaults are really triggered by a lack of liquidity, thus leading to more costly financial distress. The results also seem consistent with the view that commercial mortgage defaults occur only if the borrower does not have the cash to make

Table 6. Estimates of cumulative default probabilities by specific durations by selected mortgage and borrower characteristics (with ACLI weighting).

	Predicted by Model		
Characteristic	5 years	10 years	15 years
LTV category:			
65%	2.70	7.43	14.56
75%	2.93	8.04	15.71
85%	3.20	8.76	17.05
DCR category:			
1.15	2.56	8.06	16.92
1.25	2.27	7.16	15.11
1.50	1.67	5.31	11.33
Large	1.70	4.59	8.78
Apartment	2.62	7.03	13.28
Industrial	2.72	7.27	13.72
Office	2.58	6.91	13.07
Individual	2.93	7.82	14.72
Partnership	3.54	9.41	17.57
Corporation	2.62	7.02	13.27

Note. Differences between the default probabilities for the categories are statistically significant at the 0.10 level (or better) for all covariates except apartment, individual, partnership, and corporation.

Table 7. Effect of LTV and DCR on cumulative default probabilities by specific durations (with ACLI weighting).

	Estimated Minus Actual Default Probabilities					
	LTV Category			DCR Category		
Specific Durations	65%	75%	85%	1.15	1.25	1.50
A. 10% increase in LTV or 10% decrease in DCR						
5 years	0.14	0.20	0.27	0.38	0.37	0.34
10 years	0.39	0.53	0.72	1.17	1.14	1.05
15 years	0.72	0.99	1.33	2.32	2.29	2.14
B. 33% increase in LTV or 33% decrease in DCR						
5 years	0.54	0.77	1.07	1.50	1.47	1.38
10 years	1.45	2.04	2.81	4.51	4.47	4.26
15 years	2.71	3.77	5.13	8.72	8.76	8.55
C. 50% increase in LTV or 50% decrease in DCR						
5 years	0.91	1.31	1.86	2.57	2.51	2.49
10 years	2.43	3.47	4.87	7.66	7.69	7.56
15 years	4.51	6.36	8.78	14.50	14.75	14.87

Note. Calculations are for small, floating-rate retail property loans. Borrower type is other.

the debt payments. Much of the evidence that has been offered in support of this view has been very slight, as the methods used to estimate commercial mortgage defaults have generally been inadequate. Also, previous researchers have generally been forced to relate the default incidence on individual commercial mortgages to variables observed at the time of origination, to the age of the mortgage, and to macro-economic data, or they have been forced to relate highly aggregate measures of default incidence to national macro-economic variables. The results herein show that variables observed at the time of origination are fairly poor predictors of commercial mortgage defaults.

5. Default forecasts

Turning now to the default forecasts, Table 8 presents the errors for the population weighted and unweighted default models for each year of forecast up to 1999. The error is defined as the difference between the predicted and observed default rates, so a negative error indicates an under-prediction and a positive error indicates an over-prediction. The forecasts are calculated in the following manner. First, using data from ACLI's nationwide survey of long-term (over one year) mortgage commitments on commercial properties in

Table 8. Default estimates for commercial mortgage market.

		Model			
		Without Weights		With Weights	
Year	Actual (%)	Estimate (%)	Estimate Minus Actual (%)	Estimate (%)	Estimate Minus Actual (%)
1982	0.79	0.09	− 0.70	0.06	− 0.73
1983	0.94	0.42	− 0.52	0.27	− 0.67
1984	0.87	0.77	− 0.09	0.49	− 0.37
1985	1.06	1.47	0.41	0.95	− 0.11
1986	2.34	1.62	− 0.71	1.06	− 1.27
1987	2.90	2.29	− 0.61	1.51	− 1.39
1988	2.72	2.19	− 0.53	1.45	− 1.27
1989	2.68	1.37	− 1.31	0.90	− 1.78
1990	3.27	2.22	− 1.05	1.46	− 1.81
1991	5.47	5.87	0.40	3.92	− 1.54
1992	6.99	9.87	2.88	6.87	− 0.12
1993	6.00	9.61	3.62	7.15	1.16
1994	4.48	5.86	1.38	4.80	0.32
1995	3.21	2.32	− 0.88	2.04	− 1.16
1996	2.31	1.89	− 0.42	1.59	− 0.72
1997	1.36	1.33	− 0.02	1.04	− 0.32
1998	0.68	0.77	0.10	0.52	− 0.16
1999	0.33	0.92	0.59	0.61	0.28
MSE of estimate			1.33		0.83

the United States (including maturing balloon mortgages which have been refinanced for more than one year at current market terms), we compute a series of conditional default rate forecasts for commercial mortgage loans originated in 1974–1999. Next, unconditional default rates are then simulated, on a quarterly basis, by summing across all loan cohorts (weighted by total dollar amount outstanding). The forecasts are then compared to industry-wide default and delinquency rates in the commercial mortgage market, as reported quarterly by ACLI in their delinquency rate survey. In each and every case, we first re-scale the model's constant term in order to match ACLI's different definition of default. The effect of this is to improve the ability of the model to forecast, and lessen the importance of the heterogeneity *per se*.

One obvious feature of the results in Table 8 is that the mean-square-error (MSE) for the population-weighted forecasts is generally much smaller than that for the corresponding unweighted forecasts. The improvement in fit is particularly noticeable in the early 1990s, when actual default rates were in the 6–7 percent range. Inspection of the signs of the forecasting errors in Table 8 reveals a distinct tendency on the part of the unweighted model to over-predict during this period. In contrast, the population-weighted model is able to predict the sharp upturn in commercial mortgage default rates in the early 1990s better than the unweighted model. The population-weighted model also is able to do a reasonably good job of forecasting defaults in the late 1990s, especially when the forecasted default rate is automatically adjusted to the most recently available data in each period.

In Figure 1, the model's prediction (both weighted and unweighted) is plotted against actual default rates in the commercial mortgage market. The results suggest that option-theoretic models of commercial mortgage pricing have substantial real-world relevance. Current trends place a great deal of emphasis on understanding commercial mortgage default rates. For issuers of commercial mortgage-backed securities, assumptions related to expected mortgage default play a key role in both security structure and design. Rating

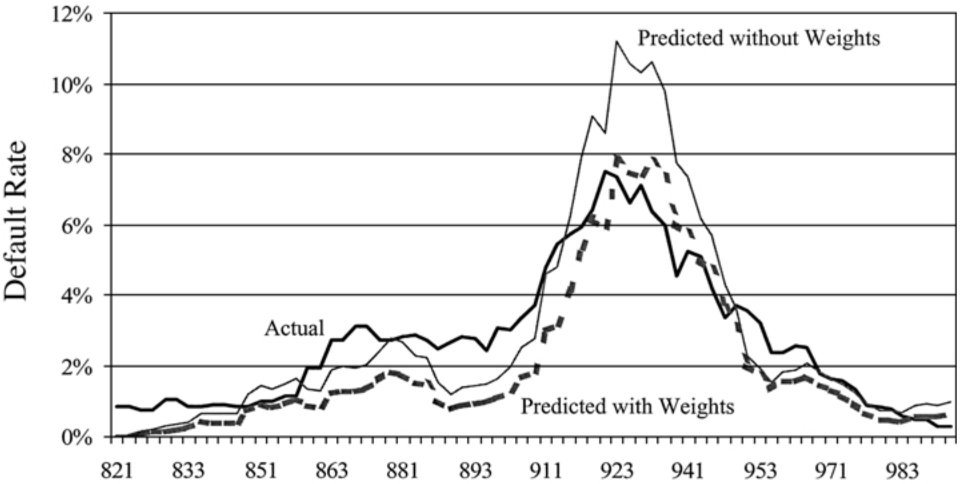

Figure 1. Actual vs. predicted default rates.

agencies are also keenly interested in the characteristics of commercial mortgage default, as security ratings are directly related to assumptions regarding both mortgage default and loss severity. These assumptions impact not only required credit enhancement, but also subordination levels associated with bond structuring. Servicers and special servicers of commercial mortgage-backed securities are intensely interested in mortgage default, as mortgage default impacts the decision making of both of these functions with regard to financially distressed loans. Finally, investors are concerned with mortgage default, as it continues to be a driving factor in the initial pricing and on-going valuation of both whole loans and commercial mortgage-backed securities. Default forecasts such as those presented here are also important for evaluating the extent of adverse selection in the market for commercial mortgage-backed securities.

6. Conclusion

This paper investigates the determinants of commercial mortgage defaults using a proportional hazards model with competing risks. Our study identifies several new facts, which raise a number of challenging questions.

First, we find that the probability of default is positively related to the contemporaneous loan-to-value ratio. Second, we find that the property's contemporaneous debt-service-coverage ratio significantly affects the probability of default, and that debt-service-coverage ratios are of more importance in explaining commercial mortgage defaults than are loan-to-value ratios. One possible argument is that default will never be necessary for a commercial mortgage borrower if the NOI is enough to cover debt payment. Also, commercial mortgage borrowers are likely to approach the default decision differently from the single-family borrower. The net equity position could still be the primary determinant of default, but the transaction costs associated with commercial mortgage default differ significantly from those of single-family properties.

Third, we find evidence to suggest that there benefits in weighting the sample estimates to fit the population. Use of a weighting process to correct for the problem of originator bias has a sound foundation in the literature.

Fourth, studying defaults in years with different economic climates, it is possible to discover how differences in interest rates, borrower type, and other loan variables affect commercial mortgage defaults. The model described in this article fits observed default rates fairly well, and does a reasonably good job in predicting turning points.

Notes

1. The approach adopted in this paper is also similar to that of Ambrose and Sanders (2000). Ambrose and Sanders specify a model of the competing risks of commercial mortgage default and prepayment, but utilize data on commercial mortgage defaults over a period in which there were very few defaults.
2. The approach undertaken here is similar to Deng et al. (2000), and Ciochetti et al. (2002). None of these studies, however, has allowed for originator bias.

3. As the DCR at origination increases, the probability of default generally decreases, *ceteris paribus*. This suggests that original DCR should have a negative coefficient. Our use of original DCR to predict commercial mortgage default is similar to of Archer et al. (2000). They find, however, that original values of DCR and LTV do not help predict commercial mortgage defaults. They attribute their findings to an endogeneity problem; that is, high-risk borrowers being forced to choose lower LTVs and higher OCRs in order to mitigate risk. Archer, Elmer, Harrison, and Ling's results are limited, however, by the lack of information on contemporaneous OCRs and LTVs.
4. For an overview of the option-theoretic pricing of mortgages, see Kau and Keenan (1995). Also, see, for example, Titman and Torous (1989), and Kau et al. (1990).
5. We have also estimated the model correcting for unobserved heterogeneity, though the correction had little effect on actual forecasting ability of the model as specified here.
6. Loan status indicators are available on a quarterly basis and include the categories of active, 30, 60, or 90 days delinquent, restructured, paid, prepaid, in process of foreclosure, or foreclosed. For purposes of this study, our interest includes prepaid loans and/or those in process of foreclosure, or foreclosed.
7. This is consistent with the burnout phenomenon as widely observed in the residential literature (see, e.g., Deng et al., 2000).

References

Ambrose, B., and A. B. Sanders. (2000). ''Commercial Mortgage-Backed Securities: Prepayment and Default,'' Ohio State University Working Paper.

Archer, W. R., P. J. Elmer, D. M. Harrison, and D. C. Ling. (2000). ''Determinants of Multifamily Mortgage Default,'' FDIC Working Paper 99–2.

Ciochetti, B. A. (1993). ''Diversification of Commercial Mortgage Porfolios,'' University of Wisconsin, Unpublished PhD Dissertation.

Ciochetti, B. A., Y. Deng, B. Gao, and R. Yao. (2002). ''The Termination of Lending Relationships through Prepayment and Default in the Commercial Mortgage Markets: A Proportional Hazard Approach with Competing Risks,'' *Real Estate Economics* 30(4), 595–633.

Ciochetti, B. A., and T. J. Riddiough. (1999). ''Timing, Loss Recovery and Economic Performance of Foreclosed Commercial Mortgages,'' University of North Carolina and MIT Working Paper.

Deng, Y., J. M. Quigley, and R. Van Order. (2000). ''Mortgage Terminations, Heterogeneity and the Exercise of Mortgage Options,'' *Econometrica* 68(2), 275–308.

Follain, J. R., and J. Ondrich. (1997). ''Ruthless Prepayment? Evidence from Multi-family Mortgages,'' *Journal of Urban Economics* 41(1), 78–101.

Han, A., and J. A. Hausman. (1990). ''Flexible Parametric Estimation of Duration and Competing Risk Models,'' *Journal of Applied Econometrics* 5(1), 1–28.

Kau, J. B., and D. C. Keenan. (1995). ''An Overview of the Option-Theoretic Pricing of Mortgages,'' *Journal of Housing Research* 6(2), 217–244.

Kau, J. B., D. C. Keenan, W. J. Muller, III, and J. Epperson. (1990). ''Pricing Commercial Mortgages and Their Mortgage-Backed Securities,'' *Journal of Real Estate Finance and Economics* 3(4), 333–356.

Pavlov, A. D. (2000). ''Competing Risks of Mortgage Termination: Who Refinances, Who Moves, and Who Defaults?'' *The Journal of Real Estate Finance and Economics* 23(2), 185–212.

Suyoshi, G. (1992). ''Semiparametric Proportional Hazards Estimation of Competing Risks Models with Time-Varying Covariates,'' *Journal of Econometrics* 51(1), 25–58.

Titman, S., and W. N. Torous. (1989). ''Valuing Commercial Mortgages: An Empirical Investigation of the Contingent Claims Approach to Pricing Risky Debt,'' *Journal of Finance* 44(2), 345–373.

Vandell, K., W. Barnes, D. Harfzell, D. Kraft, and W. Wendt. (1993). ''Commercial Mortgage Defaults: Proportional Hazards Estimation Using Individual Loan Histories,'' *AREUEA Journal* 20(4), 55–88.