The Performance of Default Risk Structural Models on Commercial Mortgages: An Empirical Investigation

Yi-Kang Liu
Pentagon Federal Credit Union

George M. Jabbour
The George Washington University

Richard K. Green
The George Washington University

Abstract

This paper uses the first-passage-time approach to estimate default probabilities of commercial mortgages and the Receiver Operating Characteristic (ROC) approach to empirically test the cash flow proposition of Vandell (1995). The focus is on comparing the performance between a single trigger model and a double-trigger model. Using 17,616 lockout commercial loans issued between 1995 and 2001, we find the property value model performs the best. In addition, the results provide a partial support to the cash flow proposition.

1. Introduction

The classical option-pricing framework developed by Black and Scholes (1973) and Merton (1974) characterizes default as a borrower's put option which will be "ruthlessly" exercised when the value of the underlying asset reaches a certain threshold. This framework has been criticized for the inconsistency between its prediction of the default and actual default behavior[1]. To reconcile the difference, Vandell (1995) posited a cash flow proposition that suggests that solvency matters: the structural models that incorporate both cash flows and loan-to-value ratio for predicting default will perform better than conventional single-trigger models that focus on LTV alone.. In this view, default is suboptimal, or even unnecessary, as long as the property can generate sufficient net operating income to cover expenses and debt service payments.

In this paper, we empirically test Vandell's (1995) cash flow proposition. Specifically, we investigate whether a double-trigger structural model outperforms a single-trigger structural model. Extant literature of commercial mortgage-backed securities (CMBS) option-based default models lacks studies on testing the cash flow proposition. We test the proposition using samples that contain a variety of commercial property types, such as hotel, office, multifamily, and industrial properties.

Our work contrasts with previous research on commercial mortgage default by applying a first-passage-time approach[2] and the receiver operating characteristic (ROC) curve approach[3] to real estate economics. The merit of the first-passage-time approach is that it can estimate default probabilities not only at maturity, but at any time before maturity. This is important for default risk evaluation since default could occur at any time after origination. The ROC curve approach enables us to examine model performance in each credit rating class and provides a statistical test on model comparison by assessing the ROC ratios.

The organization of the paper is as follows: Section two briefly reviews recent studies on commercial mortgage default. Section three introduces the first-passage-time approach and the ROC curve and explains their application to default modeling. Section four presents the data and methodologies. Section five documents the empirical results and analysis. This paper concludes with a summary of the implications of the findings.

2. Some recent studies on commercial mortgage default

Ambrose and Buttimer (2000) present an option-based model which captures the borrower's ability to reinstate the mortgage out of default prior to foreclosure. They find that loan forbearance programs, which give borrowers extra time in default, lead to both a higher delinquency rate and a postponement of default. Unfortunately, giving borrowers economic incentives to reinstate by waiving the default penalty does not appear to be effective in reducing default costs. Deng, Quigley and Van Order (2000) use a proportional hazard function to model default and prepayment as dependent competing risks to examine the termination of single-family mortgages. Their model incorporates the unobservable heterogeneity of borrowers (such as the differences in risk preference across borrowers) and estimates the unobservable heterogeneity with a baseline hazard function. Their results find that the option-based model does a "good" job of explaining default and prepayment behavior when transaction cost and borrowers' heterogeneity are excluded.

Ambrose, Capone and Deng (2001) decompose the boundary conditions for the optimal default exercise and investigate any impact of economics dynamics on the optimal default timing. Overall, their results support the option-based framework in terms of the optimal default timing. However, the high variance of their results can not lead to a conclusion that all borrowers treat default as a financial option. From the perspective of economic situations, Ambrose et al. find that borrowers tend to wait until they see the actual bottom of the real estate market, rather than to default as soon as the default option is in-the-money. Goldberg and Capone (2002) use multifamily loans to test the efficiency of a double-trigger model by applying a logistic regression. Their results show that the model with a single trigger represented by LTV (or DCR) tends to overstate (or understate) the default probabilities. However, the default probabilities estimated by the double-trigger model, which reflects the average market LTV and DCR value, performs nearly identically to the mean-value model. Archer, Elmer, Harrison and Ling (2002) argue that original LTV is not an appropriate default predicator since the lenders can simply adjust the original LTV to mitigate their risk bearing in the loan underwriting process. Their empirical results confirm their argument and find that the original LTV is endogenous to the loan original process and that the DCR is more important in explaining default behavior than original LTV.

Ciochetti, Deng, Lee, Shilling and Yao (2003) use proportional hazards models to examine the ability of structural default risk model to forecast commercial mortgage defaults. Similar to Archer, Elmer, Harrison and Ling (2002), they find that contemporaneous LTV are of less importance than contemporaneous DCR in explaining commercial mortgage defaults. Tu and Eppli (2003) define commercial mortgage default risk as the combination of the default risk during the term and the extension risk at the maturity. They use a double-trigger model (property value and property cash flow) to investigate the relation between default risk and default risk premium. Their Monte Carlo simulation results show that models based solely on property value tend to overestimate the default risk as well as the total default risk premium. However, models with double triggers underestimate the default risk but did not significantly underestimate the total default risk premium due to the high extension risk at maturity.

3. The first-passage-time approach and the ROC curve

3.1 The first-passage-time approach

Black and Cox (1976) develop the first-passage-time approach to evaluate the effect of various bond indenture provisions by broadening the scope of Merton (1974) framework. To evaluate the probability that a firm defaults, Black and Cox estimate the distribution of the firm's value at future time t ($V(t)$), conditional on the value at current time t ($V(t)$). They show that the risk neutral probability that default takes place ($V(t) < K$) at time t conditional on that the firm does not default at current time *t* is as follows:

$$P(V(t) < K \mid t > t) = N\left(\frac{\frac{\ln K_t}{\ln V_t} - (r - a - \frac{1}{2} s^2)(t - t)}{\sqrt{s^2 (t - t)}} \right) + \left(\frac{K_t}{V_t} \right)^{2(r-a-g)/s^2} N\left(\frac{\frac{\ln K_t}{\ln V_t} + (r - a - \frac{1}{2} s^2)(t - t)}{\sqrt{s^2 (t - t)}} \right) \quad (1)$$

where V_t is the value of the firm at time *t*, K_t is the bankrupt threshold which is a deterministic exponential function of time *t*, r is constant interest rate, a is the payout ratio, s is the standard deviation of the growth rate of the firm value, and g is the cost of debt.

Leland and Toft (1996) generalize Black and Cox (1976) model by deriving the first-passage-time probability density function based on a constant default barrier. Under a risk neutral world, Leland and Toft assume the value of the firm (*V*) follows a continuous diffusion process:

$$\frac{dV}{V} = (r - d)dt + s\,dz \quad (2)$$

where r is the constant risk free rate, d is the constant fraction of value paid out to security holders, s is the constant standard deviation of the growth rate of the firm value, dz is the increment of a standard Brownian motion.

The process in equation (2) continues until V falls to a constant default trigger V_B. For a bond at maturity t with constant coupon payment $c(t)$ and principal $p(t)$, Leland and Toft show that the distribution of first-passage time $F(t)$ can be expressed as follows:

$$F(t) = N[h_1(t)] + \left(\frac{V}{V_B}\right)^{-2a} N[h_2(t)] \tag{3}$$

where $N(.)$ is the cumulative standard normal distribution,

$$h_1(t) = \frac{-b - as^2 t}{s\sqrt{t}}, \qquad h_2(t) = \frac{-b + as^2 t}{s\sqrt{t}}$$

where $a = \dfrac{r - d - \dfrac{s^2}{2}}{s^2}$, $b = \ln\left(\dfrac{V}{V_B}\right)$,

r is the constant risk free rate, d is the constant fraction of value paid out to security holders, s is the constant standard deviation of the growth rate of the firm value, dz is the increment of a standard Brownian motion.

In fact, equation (3) is compatible with equation (1) except the default threshold changed from an exponential function, K_t, to a constant V_B.

3.2 Receiver operating characteristic (ROC) curve

ROC curve is a graphical plot of the cumulative rate of accurate model prediction against the cumulative rate of inaccurate model prediction for a binary classifier system (i.e. default or non-default) given by the pre-specified thresholds. This approach was initially employed in the field of electronic signals and diagnostic (see Green and Swets (1966), Hanley (1989), and Pepe (2000)). Until recently, academics and practitioners in finance successfully apply the ROC approach to empirical test on the performance of default risk models (Stein (2002), Kealhofer (2003), and Engelmann et al. (2003)).

The most important component of the ROC curve is the contingency table. Contingency table documents the model outcomes compared to the actual results across all possible cutoff points (thresholds). Table 1 shows the simplest form of contingency table with only one cutoff point which divides the model outcomes into two groups: high default probabilities and low default probabilities. The number of cutoff points is arbitrary but highly depends on the sample size. In general, with a large sample size, it is reasonable to use more than one cutoff point to investigate the model performance.

There are four steps to construct the ROC curve. First, estimate the default probability of each loan and order the loans from the highest default probability to the lowest default probability. Second, divide all loans into two groups (projected defaults and projected non-defaults) based on the cut off point. Third, break down the projected defaults and projected non-defaults into four categories (hit, type I error, type II error, and correct rejection) by comparing the predictions to the actual outcomes. Fourth, plot the cumulated false alarm rate (defined as type II error /(type II error + correct rejection)) along the x axis against the corresponding cumulated hit rate (defined as hit /(hit + type I error)) along the y axis. Figure 2 shows the hypothetic ROC curve associated with the contingency table in Table 1.

For a good model, the false alarm rate should be low and the hit rate should be high. Consequently, its ROC curve lies above the random model represented by the 45 degree line starting from point zero. In fact, the more the area under the ROC curve is (area OEFG in the Figure 2), the more accurate the default prediction is.

4. Data and methodology

4.1 Data

Data are the conduit loans collected from the Trepp CMBS Deal Library. The origination years of samples are from 1995 to 2001 and the observation date is September 2004. The reason for this particular time horizon is twofold: first, there are a small number of loans observable in 2004 that were originated before 1995, because most commercial loans bundled into CMBS have terms of less than ten years. Second, the peak years for CMBS default are the third to the seventh year after origination[4]. The loans in our sample are between three and nine years old.

In addition, to exclude the impact of prepayment and credit enhancement on default, the loans in our sample are limited to locked-out fixed-rate loans without cross collateralization. The set of data includes 17,161 commercial mortgage loans that are backed by a variety of property types: hotel (966 loans), industrial (1,516 loans), multi-family (5,094 loans), office (3,084 loans), retail (4,928 loans) and others (1,573 loans). Geographically, the highest number of loans comes from the South (6,215 loans). The Midwest has the fewest loans (2,405). In total, there are 327 defaulting loans[5] and 16,834 non-defaulting loans. Table 2 presents descriptive statistics and geographic locations of the samples.

4.2 Methodology

The purpose of this study is to investigate whether the double-trigger proposition can reconcile the discrepancy between the actual and expected default behavior in commercial mortgages. To this end, we first build three different models (property value model, cash flow model and double-trigger model) characterized by default triggers and then compare the model performance. To alleviate the impact of ruthless default assumption, we assume transaction costs to be 2% of the original property value. In addition, for the double-trigger model, we run two variants: one where we assume the correlation between the growth rate of the property value and the property net operating income to be either 0.5, another where we assume 0.8[6].

4.2.1 Property value model

The value of a property is assumed to follow a geometric Brownian motion:

$$\frac{dV_t}{V_t} = m_V dt + s_V dW_t^V \tag{4}$$

where V_t is the value of the property at time t, m_V is the mean growth rate of return on property, s_V is the volatility of rate of return on property, W_t^V is Wiener process associated with property at time t.

The input parameters (m_V and s_V) are determined as follows. Under the risk-neutral probability measure, the drift term m_V is replaced by the risk-free rate which can be estimated by averaging the one-year Treasury constant maturity rates from 1995 to 2001[7]. In addition, we use the standard deviations of the NCREIF unsmoothed returns on different property types, estimated by Georgiev, Gupta and Kunkel (2003) during the period between 1990 and 2001, as the proxies of the volatilities of rate of return on properties (s_V). The advantage of the unsmoothed returns is to avoid the potential bias of the appraisal-based NCREIF index return due to the autocorrelation problem (Gatzlaff and Geltner (1998) and Clayton, Geltner and Hamilton (2001)).

By definition, default occurs when the value of the property at time t (V_t) falls below the mortgage balance (L, decreasing with time). The cumulative probability of default at time t can be estimated from equation (3). Recognizing that the initial LTV is endogenous to the loan underwriting process as suggested by Archer, Elmer, Harrison and Ling (2002), we model contemporaneous LTV as the default predictor for the property value model. Specifically, we estimate multiple default probabilities based on the contemporaneous LTV at different time horizons between current time and loan origination and then calculate the mean of the multiple default probabilities as the expected default probability for a commercial mortgage loan.

4.2.2 Cash flow model

The value of net operating income[8] of a property is assumed to follow a geometric Brownian motion:

$$\frac{dN_t}{N_t} = m_N dt + s_N dW_t^N \qquad (5)$$

where N_t is the value of the net operating income at time t, m_N is the mean growth rate of net operating income, s_N is the volatility of the growth rate of net operating income, W_t^N is Wiener process associated with net operating income at time t.

Following Tu and Eppli (2002), we assume the mean growth rate of net operating income (m_N) to be constant at 3%, which is a reasonable estimate based on economic conditions during the sample period. For each property type, the volatility of the net operating income (s_N) is approximated by the volatility of real estate investment trust (REIT) with the similar property type. Ambrose, Highfield and Linneman (2005) show that the implied capitalization rate (ICR), defined as the net operating income of REIT divided by the average real estate value, is an appropriate measure for the REIT growth prospects. Since the growth rate of REIT is highly correlated with the growth rate of the net operating income, we use the standard deviations of ICR estimated by Ambrose, Highfield and Linneman (2005) during the sample period from 1990 to 2001 as the proxy for parameter s_N.

Unlike the property value model which has a time-decreasing threshold (*L*, mortgage balance), cash flow model has a constant default threshold (*D*, monthly amortized mortgage payment). Thus, the expected default probabilities of commercial mortgage loans at origination can be obtained by assessing contemporaneous DCR based on equation (5) and (3).

4.2.3 Double-trigger model

Owing to the lack of a closed-from solution, we analytically estimate the default probability of the double-trigger model. Following Goldberg and Capone (2002), the default probability of the double-trigger model is defined as:

$$prob(\ln V < \ln L \cap \ln N < \ln D) \tag{6}$$

where *V* is the property value. *L* is the mortgage balance which is decreasing with time. *N* is the monthly net operating income. *D* is the constant monthly mortgage payment. Consistent with the assumptions of property value model and cash flow model, *V* and *N* follow the dynamic in equation (4) and (5) respectively. In addition, *V* and *N* are assumed to be jointly lognormally distributed. By standardizing the normal distributed random variable ln *V* and ln *N*, the default probability (6) becomes:

$$prob\left(\frac{\ln V - E(\ln V)}{std(\ln V)} < \frac{\ln L - E(\ln V)}{std(\ln V)} \bigcap \frac{\ln N - E(\ln N)}{std(\ln N)} < \frac{\ln D - E(\ln N)}{std(\ln N)}\right)$$

$$= prob(x < a \bigcap y < b) \qquad (7)$$

where *x* and *y* are standard bivariate normal.

The cumulative probability of the bivariate normal distribution *x* and *y* in equation (7) is analytically approximated by Drezner (1978) approach. In particular, we estimate multiple default probabilities at various time periods and calculate the mean of these probabilities as the estimated default probability for the double-trigger model.

4.3 ROC accuracy ratio and empirical test

The model performance of default prediction is measured by the ROC curve and the ROC accuracy ratio estimated by the Mann-Whitney *U* statistic of each model[9]. The higher the *U* statistic is, the more accurate the model is.

Following Engelmann, Hayden and Tasche (2003), we calculate a statistic (*T*) to empirically test the difference between the two models:

$$T = \frac{(\hat{U}_i - \hat{U}_j)^2}{s^2_{\hat{U}_i} + s^2_{\hat{U}_j} - 2s_{\hat{U}_i,\hat{U}_j}} \qquad (8)$$

Where $\hat{U}_i$ and $\hat{U}_j$ are the Mann-Whitney *U* statistics of model *i* and model *j* respectively. $s^2_{\hat{U}_i}$ and $s^2_{\hat{U}_j}$ are the variances of Mann-Whitney *U* statistic of model *i* and model *j* respectively. $s_{\hat{U}_i\hat{U}_j}$ is the empirical covariance of Mann-Whitney *U* statistics between model *i* and *j*.

Engelmann, Hayden and Tasche (2003) point out that the T statistic in equation (8) is asymptotically chi-square distributed with one degree of freedom. The larger the T statistic indicates the more significance that the model i outperforms model j. In other words, given a significant level a and the variances of U statistics[10], we can empirically test the cash flow proposition by simply examining the significance of T statistics.

4.4 Example

In Appendix, we provide a numerical example to demonstrate how to estimate the default probability, construct the ROC curve and calculate the ROC accuracy ratio in details.

5. Results

Figure 2 reports the ROC curves and the ROC accuracy ratios. The ROC curve of the property value model is always above the ROC curve of a random model. This result indicates that the default prediction of the property value model is more accurate than a random guess. In addition, the ROC accuracy ratio of the property value model is 61.7625%, the highest among all models. This result provides initial evidence implying the advantage of LTV as a default predicator. However, despite the high ROC accuracy ratio of the property value model, a rigorous statistic test must be conducted to justify the conjecture that LTV is the predominant factor in explaining default behavior. We will address this part later in this paper.

The shape of the ROC curve for the cash flow model differs from the property value model. For cash flow model, the ROC curve tends to hump when the false alarm rate is between 0 and 0.1. Since the ROC curve is developed by matching the hit rate against the false alarm rate for each rated class, the evidence that the ROC curves of cash flow model hump in the very beginning indicates that the cash flow model performs better among the lowest quality loans than it does among the others. Nonetheless, the ROC ratio of the cash flow model is 55.2304%, about 6% lower than the property value model.

For double-trigger model, the shapes of the ROC curves are similar to the ROC curve of the property value model and have a generally upward sloping trend. From the perspective of prediction performance, the double-trigger model yields the ROC ratios of 60.3322% (when the default trigger correlation is 0.5) and 60.7277% (when the default trigger correlation is 0.8). The evidence of ROC ratio increasing with default trigger correlation may attribute to the dominant effect of LTV, which has more impact on the expected default behavior than DCS when LTV and DCS are highly correlated.

Table 3 reports the model performance between the single trigger model and the double-trigger model. The empirical results are summarized as follows.

5.1 Property value model vs. cash flow model

The chi-square test shows that the performance of the property value model is significantly different from that of the cash flow model. The confidence coefficient is significant at the 99% level. Recall from Figures 2, the ROC ratio of property value model (61.7625%) is larger than that of the cash flow model (55.2304%). Taking both empirical results together, we conclude that the property value model outperforms the cash flow model. In other words, the commercial mortgage default prediction estimated by LTV performs better than the default prediction estimated by DCR.

5.2 Property value model vs. double-trigger model

Unlike the previous results, the chi-square test shows that the property value model does not significantly differ from the double-trigger model. Although the

ROC ratio of the property value model (61.7625%) is larger than that of the double-trigger model (60.3322% and 60.7277%), the statistic from Table 3 is insufficient to justify the superiority based on model performance. In other words, the combined effect of LTV and DCR on expected default behavior does not significantly differ from that of a single LTV factor, or vice versa. As such, this result leads to a rejection of hypothesis test on cash flow proposition which conjectures differential performance between property value model and double-trigger model.

5.3 Double-trigger model vs. cash flow model

The chi-square test in Table 3 indicates that the cash flow model is statistically significantly different from the double-trigger model. The confidence coefficient is significant at the 95% level. Since the ROC ratios of double-trigger model (60.3322% and 60.7277%) are larger than that of the cash flow model (55.2304%), we can conclude that the double-trigger model outperforms the cash flow model. That is, the commercial mortgage default prediction estimated by LTV and DCR is more accurate than the default prediction estimated by DCR alone. Contrary to the performance between property value model and double-trigger model, this empirical evidence is consistent with the cash flow proposition.

5.4 Low-quality loans

One of the interesting findings in this study is the model predicting power for low-quality loans. We define low-quality loans as the loans with default probabilities that are at the top 12% of the loan distribution. A model with better discriminative power should demonstrate a high hit rate on this class of loans. Among the three structural models, we find that, although the cash flow model has the lowest ROC accuracy ratio for the full sample of loans, it provides relatively strong explanatory power to low-quality loans. Our results show that the cumulative hit rate at 12% threshold for cash flow model is 22.02%, which is better than 18.04% for property value model and 20.49% for double-trigger model with a default trigger correlation of 0.5, or 19.88% for double-trigger model with a default trigger correlation of 0.8. For mortgage underwriters, this empirical evidence suggests that while dealing with low-quality loans, liquidity risk measured by DCR, rather than LTV, should be the key factor affecting borrowers' default decisions.

6. Summary and conclusions

Based on the option-pricing framework, this study provides a theoretical approach to estimate default probabilities of commercial mortgages and an empirical test on the relation between default triggers and default behavior. Using 17,616 lockout commercial loans issued between 1995 and 2001, we empirically test the cash flow proposition by comparing the ROC accuracy ratios. Our results show that, the ROC accuracy ratios of all structural models are larger than 50%, the benchmark of a random model. This evidence suggests that the structural models do have better model performance than a random model in detecting defaults. In addition, the ROC accuracy ratio of the property value model is larger than those of the cash flow model and the double-trigger model. This evidence implies that the LTV ratio, the underlying default trigger for the property value model, is the most dominant factor affecting the default behavior of commercial mortgages.

For model comparison, we find that both the property value model and the double-trigger model outperform the cash flow model. However, the default prediction accuracy between the property value model and the double-trigger model is not significantly different. Therefore, the superiority between the property value model and the double-trigger model can not be specifically inferred from our empirical results. Namely, the cash flow hypothesis is not fully supported by the outcomes of structural models.

The lack of support to cash flow proposition could be attributed to the unique characteristic of DCR. Unlike LTV which is important to default behavior throughout the entire duration of CMBS, DCR is critical to default decision only while it is close to, or even less than one. In this case, the borrowers are exposed to liquidity risk and are more likely to default ruthlessly. On the other hand, if the DCR is at a certain level which the property can generate positive cash flow, the impact of DCR on default behavior diminishes simply because the borrowers are solvent and the possibility of ruthless default due to illiquidity is slim. As a consequence, LTV takes over and becomes the dominant factor on default behavior. This conjecture is consistent with the argument in Ciochetti, Deng, Lee, Shilling and Yao (2003) who argue that default is unnecessary for commercial mortgage if the net operating income is sufficient to cover debt payment.

One manifestation of our conjecture is the evidence that the cash flow model exceeds both property value model and double-trigger model in detecting defaults among low quality loans and that the performance of cash flow model diminishes as the quality of the loans increases. In Figure 2, the ROC curve of cash flow model outperforms other models in terms of default hit rate when the loans are at the top 12% of the expected default probabilities estimated by DCR. After this threshold, the efficiency of the cash flow model begins to decrease. While the quality of loans increases by moving from top 12% to top 20% of the expected default probabilities, both double-trigger model and property value model outperform the cash flow model by showing a better hit rate in default prediction.

In conclusion, this study investigates the validity of structural models and tests the cash flow proposition. The results show that the structural models do have better discriminative power than a random model. In addition, the evidence that the property value model has the highest ROC accuracy ratio implies that the LTV ratio is the major factor influencing the occurrence of default. On the other hand, the cash flow proposition is only partly supported by the empirical findings.

Despite the fact that the double-trigger model significantly outperforms the cash flow model, the evidence that the insignificant difference between the model performance of the double-trigger model and the property value model does not support the cash flow proposition.

References

Ambrose, B. W., and R. J. Buttimer. 2000. Embedded Options in the Mortgage Contract. Journal of Real Estate Finance and Economics 21 (2): 95-111.

Ambrose, B. W., C. A. Capone, and Y. Deng. 2001. Optional Put Exercise: An Empirical Examination of Conditions for Mortgage Foreclosure, Journal of real Estate Finance and Economics 23 (2): 213-234.

Ambrose, B. W., M. Highfield, and D. Linneman. 2005. Real estate and economies of scale: the case of REITs. Real Estate Economics 33 (2): 323-350.

Ambrose, B. W., and A. B. Sanders. 2003. Commercial Mortgage-Backed Securities: Prepayment and Default. Journal of Real Estate Finance and Economics 26: 179-196.

Archer, W. R., P. J. Elmer, D. M. Harrison and D. C. Ling. 2002. Determinants of multifamily mortgage default, Real Estate Economics 30 (3), 445-473.

Bamber, D. 1975. The Area above the Ordinal Dominance Graph and the Area below the Receiver Operating Graph, Journal of Mathematical Psychology 12: 387- 415.

Black, F. and J. C. Cox. 1976. Valuing Corporate Securities: Some Effects of Bond Indenture Provisions. Journal of finance 31: 351-367.

Black, F. and M. Scholes. 1973. The Pricing of Options and Corporate Liabilities, Journal of Political Economy 81: 637-659.

Ciochetti, B. A., Y. Deng, G. Lee, J. D. Shilling and R. Yao. 2003. A Proportional Hazards Model of Commercial Mortgage Default with Originator Bias. Journal of Real Estate Finance and Economics 27: 5-23.

Clayton, J., D. Geltner and S. W. Hamilton. 2001. Smoothing in Commercial Property Valuations: Evidence from Individual Appraisals. Real Estate Economics 29 (3): 337-360.

Deng, Y., J. M. Quigley and R. Van Order. 2000. Mortgage Terminations, Heterogeneity and the Exercise of Mortgage Options. Econometrica 68: 275-307.

Engelmann, B., E. Hayden, and D. Tasche. 2003. Measuring the discriminative power of rating systems. Deutsche Bundesbank.

Gatzlaff, D. and D. Geltner. 1998. A Transaction-Based Index of Commercial Property and Its Comparison to the NCREIF Index. Real Estate Finance 15 (1): 7-22

Georgiev, G., B. Gupta and T. Kunkel. 2003. Benefits of Real Estate Investment. Journal of Portfolio Management, Special Real Estate Issue; 28-34

Goldberg, L. and C. A. Capone. 2002. A Dynamic Double-Trigger Model of Multifamily Mortgage Default, Real Estate Economics 30: 85-113.

Green, D. and J. Swets. 1966. Signal Detection Theory and Psychophysics. Peninsula Publishing: Los Altos, CA.

Hanley, J. 1989. Receiver Operating Characteristic (ROC) Methodology: The State of the Art. Critical Reviews in Diagnostic Imaging 29: 307-335.

Leland, H. E. and K. B. Toft. 1996. Optimal Capital Structure, Endogenous Bankruptcy, and the Term Structure of Credit Spreads. Journal of Finance 51: 987-1019.

Merton, R. 1974. On the Pricing of Corporate Debt: The Risk Structure of Interest Rates. Journal of Finance 29: 449-470.

Pepe, M.S. 2000. Receiver Operating Characteristic Methodology. Journal of the American Statistical Association 95: 308-311.

Tu, C. C. and M. J. Eppli. 2003. Term Default, Balloon Risk, and Credit Risk in Commercial Mortgages. Journal of Fixed Income 13: 42-53.

Vandell K. D., W. Barnes, D. Hartzell, D. Kraft, and W. Wendt. 1993. Commercial Mortgage Default: Proportional Hazards Estimation Using Individual Loan Histories. Journal of the American Real Estate and Urban Economics Association 21: 451-480.

Vandell, K. D. 1995. How Ruthless is Mortgage Default? A Review and Synthesis of the Evidence. Journal of Housing Research 6: 245-265.

Van Deventer, D. and J. Outram. (2002). The new capital accord and internal bank ratings. in Credit ratings: methodologies, rationale and default risk, Ong, M. editor, Risk book, 2002.

Appendix

For illustrative purpose, we implement the property value model to derive the default probability based on the following assumptions:

L (current loan amount) = 13,250,000

V (property value) = 15,500,000

TC (transaction cost) = 2% of property value

K (default threshold) = *L-TC*V* = 12,940,000

r (risk-free rate) = 5%

s (volatility of the growth rate of property value) = 10%

t (age of the loan) = 5 years.

Following Black and Cox (1976) and Leland and Toft (1996), the estimated default probability at current time *t* (*P(t)*) given by the fact that the loan did not default before is:

$$P(t) = N(X) + (\frac{V}{K})^{-2a} N(Y)$$

$$X = \frac{-b - as^2 t}{s\sqrt{t}}, Y = \frac{-b + as^2 t}{s\sqrt{t}}, a = \frac{r - \frac{s^2}{2}}{s^2}, b = \ln(\frac{V}{K})$$

where $N(.)$ is the cumulative standard normal distribution. By plugging the parameters in the assumptions, we can estimate default probability as follows:

$$a = \frac{r - \frac{s^2}{2}}{s^2} = \frac{0.05 - \frac{0.1^2}{2}}{0.1^2} = 4.5, \qquad b = \ln(\frac{V}{K}) = \ln(\frac{15,500,000}{12,940,000}) = 0.1805$$

$$X = \frac{-b - as^2 t}{s\sqrt{t}} = \frac{-0.1805 - 4.5 * 0.1^2 * 5}{0.1 * \sqrt{5}} = -1.8135$$

$$Y = \frac{-b + as^2 t}{s\sqrt{t}} = \frac{-0.1805 + 4.5 * 0.1^2 * 5}{0.1 * \sqrt{5}} = 0.1990$$

$$P(t) = N(X) + (\frac{V}{K})^{-2a} N(Y) = N(-1.8135) + (\frac{15,500,000}{12,940,000})^{-2*4.5} * N(0.1990)$$

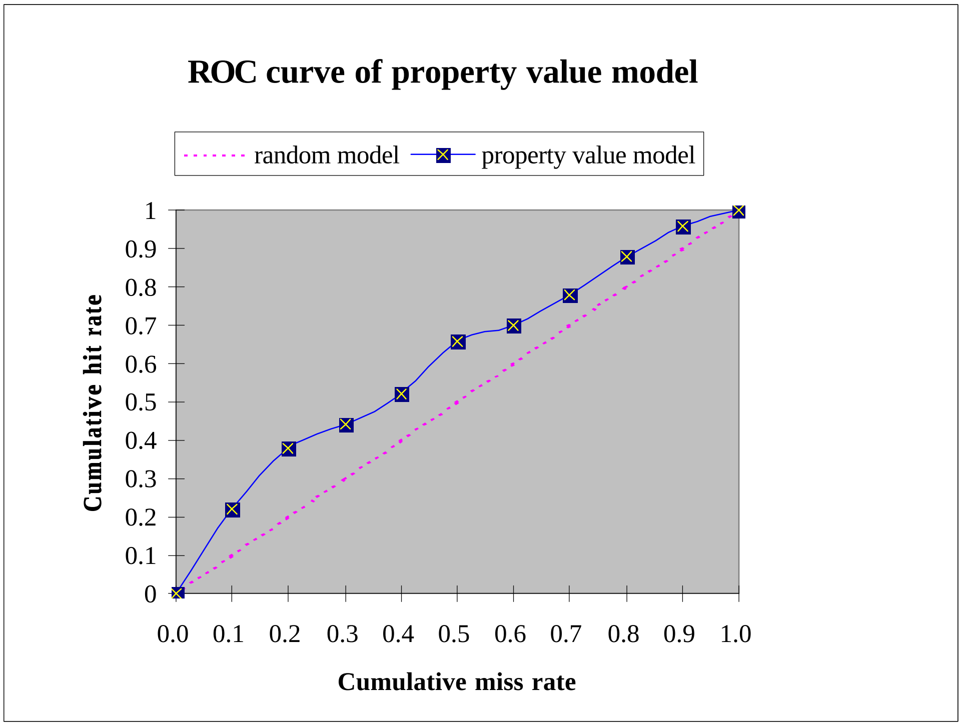
$$= 0.0348 + 0.1970 * 0.5789 = 0.1488$$

According to the calculation, the estimated default probability of this loan at age of 5 years is about 14.9%.

To develop the ROC curve, we estimate the default probabilities for all loans in the data and then sort the loans in a descending order based on the estimated default probabilities. The next step is to find the cumulative hit rate and cumulative miss rate in each decile. The following table presents the actual results of property value model by using the loans from the West.

Decile	Default	Non-default	Cumulative hit rate	Cumulative miss rate
10%	11	522	0.2200	0.0988
20%	8	525	0.3800	0.1981
30%	3	530	0.4400	0.2984
40%	4	529	0.5200	0.3985
50%	7	526	0.6600	0.4980
60%	2	531	0.7000	0.5985
70%	4	529	0.7800	0.6986
80%	5	528	0.8800	0.7985
90%	4	529	0.9600	0.8986
100%	2	536	1.0000	1.0000
Total	50	5285		

For each decile, the cumulative hit/miss rate is defined as the sum of the default/non-default loans above the decile divided by the total number of default/non-default loans. For instance, the cumulative hit rate for 30% decile is equal to (11+8+3)/50=0.44. The ROC curve (as shown in the following graph) is constructed by depicting the relation between the cumulative hit rates against the cumulative miss rates along with deciles.



Finally, we provide a simple example to illustrate how to calculate the ROC accuracy ratio (i.e. Mann-Whitney *U* statistic) for a dataset with eight hypothetical observations as following:

Prob of defaulting loans \ Prob. of non-defaulting loans	15%	10%	8%	3%	0.5%
13%	0	1	1	1	1
10%	0	0.5	1	1	1
5%	0	0	0	1	1

In the above table, the eight observations include three defaulting loans (listed in the first column, percentage is the default probability of each defaulting loan) and five non-defaulting loans (listed in the first row, percentage is the default probability of each non-defaulting loan). There are total fifteen pairs (3x5) derived from the eight observations. For each pair, the score will be determined by comparing the default probabilities between defaulting and non-defaulting loans: one/zero point is given if the probability of defaulting/non-defaulting loan is larger than the probability of the non-defaulting/defaulting loan. Half point represents a tie. According to the above table, the sum of all points is 9.5. As a result, the ROC accuracy ratio is 9.5/15= 0.6333.

Footnotes

1. Vandell (1992) finds that, for commercial mortgage loans, only 8 percent of the expected defaulting loans turn out to actually default.

2. The term "first-passage-time" refers to the implicit assumption that default occurs only at the very first time when the value of the underlying asset reaches the pre-specified threshold.

3. We will discuss ROC approach in details later in this paper.

4. This empirical evidence is found by Esaki (2002).

5. Following Snyderman (1994) and Archer et al. (2002), we define default as the loans that fail to pay their monthly mortgage payments at least 90 days past due.

6. Theoretically, the value of the commercial property is positively related to the stream of projected cash flow. As a result, we expect high correlation between the growth rate of net operating income and the growth rate of the property value. Unfortunately, we don't have good information on the size of the correlation, so we run our model using to different assumptions about correlations.

7. One-year Treasury constant maturity rates are obtained from the database at the Federal Reserve Bank of St. Louis.

8. The net operating income (NOI) is defined as the total monthly revenues less operating expenses before taxes, depreciation and mortgage payments. NOI is commonly regarded as a solvency indicator for debt obligations.

9. For more details on ROC accuracy ratio estimation and Mann-Whitney *U*

statistic, see Van Deventer and Outram (2002).

10. The variances of *U* statistics are estimated by following Bamber (1975) approach.

Table 1: Contingency table

Table 1 presents a contingency table between the model predictions and the actual outcomes. The cut off point for the projected defaults and the projected non-defaults is determined by an arbitrary pre-specified threshold (default probability). The total sample size is the sum of the number of (hit) + (type II error) + (type I error) + (correct rejection).

		Actual outcomes	
		Default	Non-default
Credit ratings	Projected defaults (high default probabilities)	correct prediction (hit)	wrong prediction (type II error)
	Projected non-defaults (low default probabilities)	wrong prediction (type I error)	correct prediction (correct rejection)

Table 2: Data descriptive statistics and geographic locations

The data set includes 17,161 commercial mortgage loans with diversified property types: hotel (966 loans), industrial (1,516 loans), multi-family (5,094 loans), office (3,084 loans), retail (4,928 loans) and others (1,573 loans). There are 327 defaulting loans and 16,834 non-defaulting loans.

Panel A: descriptive statistics

	Mean	Std
Original balance	$6,505,632	$11,537,169
Current balance	$6,112,352	$10,848,255
Original LTV (%)	69.50	11.14
Original DCR	1.43	0.82
Term (month)	133.45	38.82

Source: 1995~2001 public conduit CMBS loans, Trepp CMBS Deal Library.

Panel B: geographic locations

Geographic region*	Number of total loans	Number of defaulting loans	Default rate**
Northeast	3,206	48	1.4971%
Midwest	2,405	67	2.7858%
South	6,215	162	2.6065%
West	5,335	50	0.9372%
Total	17,161	327	1.9054%

* Following the definitions by U.S. Census Bureau, Northeast includes 9 states (CT, ME, MA, NH, NJ, NY, PA, RI, and VT). Midwest includes 12 states (IL, IN, IA, KS, MI, MN, MO, NE, ND, OH, SD, and WI). South includes 16 states (AL, AR, DE, FL, GA, KY, LA, MD, MS, NC, OK, SC, TN, TX, VA and WV) and Washington D.C. West includes 13 states (AK, AZ, CA, CO, HI, ID, MT, NV, NM, OR, UT, WA and WY).

** Default rate is defined as the number of defaulting loans divided by the number of total loans in the corresponding region.

Table 3: Comparison of ROC accuracy ratios between models

Table 3 compares the ROC accuracy ratios between models. The statistic (*T*) is calculated as follows:

$$T = \frac{(\hat{U}_1 - \hat{U}_2)^2}{\mathrm{var}(\hat{U}_1) + \mathrm{var}(\hat{U}_2) - 2\,\mathrm{cov}(\hat{U}_1, \hat{U}_2)}$$

Where $\hat{U}_i$ is the Mann-Whitney *U* statistic of model *i*. $\mathrm{var}(\hat{U}_i)$ is the variance of Mann-Whitney *U* statistic of model *i*, and $\mathrm{cov}(\hat{U}_i, \hat{U}_j)$ is the empirical covariance of Mann-Whitney *U* statistics between model *i* and *j*. The *T*-statistic is asymptotically Chi-square distributed with one degree of freedom. PV is the property value model. CF is the cash flow model. DT (0.5) and DT (0.8) are the double-trigger models with the correlation coefficient of default triggers 0.5 or 0.8 respectively.

Models	T-statistic	p-value	Confidence coefficient
PV vs CF	8.29391	0.00398	0.99602*
PV vs DT (0.5)	0.42311	0.51539	0.48461
PV vs DT (0.8)	0.23274	0.62953	0.37047
DT (0.5) vs CF	4.58069	0.03233	0.96767*
DT (0.8) vs CF	5.55288	0.01845	0.98155*

* Significant at 95% level

Figure 1: Receiver operating characteristics curve

This ROC curve is constructed based on the contingency table in Table 1.

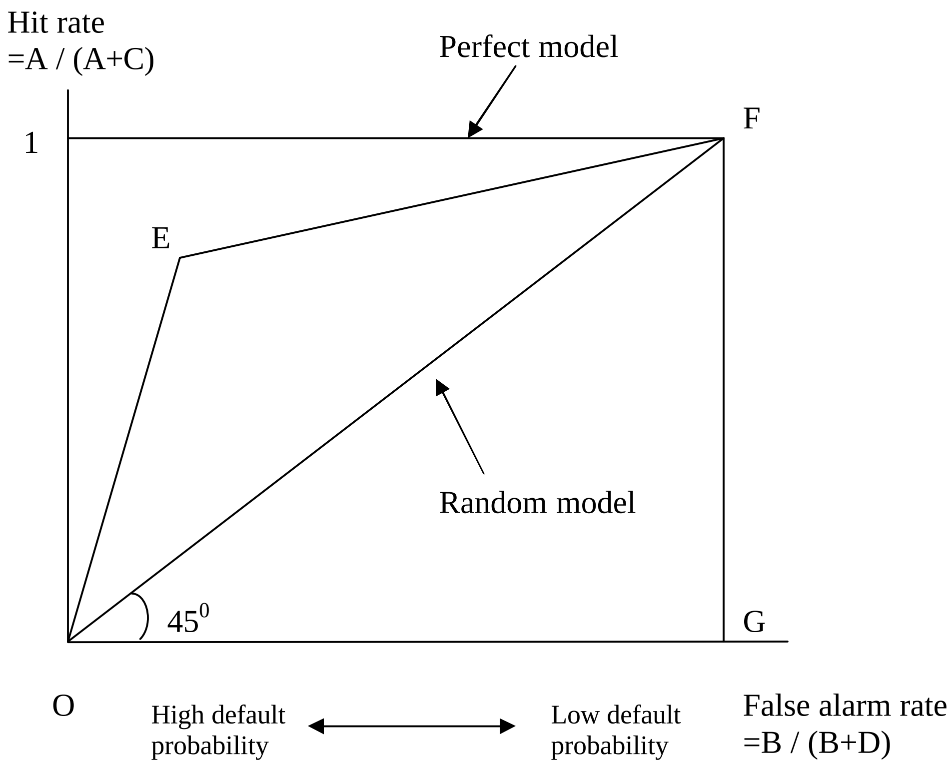


Figure 2: ROC curves and ROC accuracy ratios

Figure 2 reports the ROC curves and ROC accuracy ratios for all models. PV stands for the property value model. CF stands for the cash flow model. DT represents the double trigger model. All ROC curves are subject to 2% transaction cost.