Saunders & Allen (2010)

CHAPTER 10

Stress Testing Credit Risk Models: Algorithmics Mark-to-Future

INTRODUCTION

A key issue for bankers and regulators is internal model validation and predictive accuracy. In the context of market models, this issue has led to numerous efforts to back-test models to ascertain their predictive accuracy. The second pillar of the Basel II capital accords states that bank regulators must evaluate how well banks are assessing their capital needs relative to their risk, thereby requiring bank examiners to validate the accuracy of bank risk measurement models. Currently, under the Basel market risk-based capital requirements, a bank must back-test its internal market model over a minimum of 250 past days if it is used for capital requirement calculations. If the forecast VAR errors on those 250 days are too large, implying that the risk is underestimated on too many days, a system of penalties is imposed by regulators to create incentives for bankers to get their models right.[1]

Many observers, however, have argued that back-testing over 250 days is simply not enough, given the high standard errors that are likely to occur if the period is not representative of true market conditions. To reduce errors of this type, one suggestion has been to increase the number of past daily observations over which a back-test of a model is conducted. For example, using at least 1,000 past daily observations is commonly considered to be adequate to ensure that the period chosen is representative in terms of testing the predictive accuracy of any given model.[2] Unfortunately, even for traded financial assets such as currencies, a period of 1,000 past days requires going back in time over four years and may involve covering a wide and unrepresentative range of forex regimes.

In response to these criticisms, bank regulators conducted a stress test of the 19 largest U.S. banks (each with year-end assets exceeding $100 billion as of December 2008) that was forward looking. The test required the banks "to project their credit losses and revenues for the two years 2009 and 2010, including the level of reserves that would be needed at the end of 2010 to cover expected losses in 2011, under two alternative economic scenarios" (Board of Governors of the Federal Reserve System, April 2009). In this chapter, we discuss the design and results of the Federal Reserve stress tests of major banks in 2009.

BACK-TESTING CREDIT RISK MODELS

To appropriately back-test or stress test market-risk models, 250 observations may be regarded as too few, but it is unlikely that a bank would be able to generate anywhere near that many past time-series observations for back-testing its internal credit-risk models. For example, with annual observations (which are the most likely to be available), a bank might be able to generate only 40 past observations that cover five or six credit cycles.[3] A banker or regulator is then severely hampered from performing time-series back-testing similar to that currently available for market risk models.[4]

Even when available for back-testing of credit risk models, loan databases are often subject to substantial error in classifications. In order to compute the loss distribution for a loan portfolio, individual loans must be classified according to their default probabilities. Carey and Hrycay (2001) compare three methodologies to accomplish this:

1. The internal ratings method.
2. Mapping to external ratings.
3. Credit scoring (see Chapter 6).

These methodologies have biases that may undermine the accuracy of the estimated loss distribution. For example, the internal ratings method may be unstable if ratings criteria have changed over time or if there are insufficient data to estimate a time-invariant historical average default rate for each internal rating classification. In contrast, the efficacy of the external ratings mapping method is undermined by possible judgmental biases in assigning each individual loan to a particular external ratings classification. Finally, credit scoring models suffer from biases in model estimates that are exacerbated across different credit cycles. Carey and Hrycay (2001) find that the classification model does well in quantifying rating grades, but correctly identifies only one third of defaulting firms. Moreover, the biases introduced by errors in classification differ for investment-grade as opposed to non-investment-grade debt instruments. Some, but not all, of these problems can be alleviated if long panels of loan data are collected.[5]

Benchmarking and Accuracy Ratios

Traditionally, back-testing approaches employed by regulators have been to evaluate or stress test a given bank's model by comparing that bank's loan rating system with that of similar size banks. For example, suppose that IBM has drawn-down loans from 10 banks, each of which has an internal rating system. As discussed in more detail in the Appendix to Chapter 13, bank internal rating systems rate the credit risk of the borrower on a numerical scale (e.g., 1 to 10, where 1 is the best credit and 10 is the worst). The internal rating is based on the bank's assessment, supported by historical data, of either the probability of default, PD (unitary scale), or both PD and the loss given default, LGD (dual scale). A regulator can stress test the accuracy of the bank's internal rating process by comparing the rating of a particular company, say IBM, using 10 different banks' internal ratings (normalized to the same scale). In this way, an extreme outlier can be identified and the bank regulator can then penalize the bank by rejecting the accuracy of the bank's internal ratings model. This may cause the bank to lose the ability to use its internal ratings to calculate its capital for credit risk capital requirements, as specified in the Basel II Capital Accord (see Chapter 13). This will also increase the penalized bank's capital requirement.

The accuracy of internal ratings models is difficult to assess, however, because of the potential for either highly rated borrowers to default or borrowers with low ratings to repay their loans. That is, there is an overlap between the defaulting and nondefaulting internal ratings, as shown in Figure 10.1.

A banker may choose the cut-off point C so that all loan applicants with scores below C will be denied loans because they are expected to default, and all applicants with scores above C will be granted loans. However, Figure 10.1 shows that some successful loan applicants will default, whereas some denied loans (the cross-hatched area) would not have defaulted. The question is to determine the correct cut-off point C that will maximize the hit rate while minimizing the false alarm rate. The hit rate for a cut-off point C, HR(C) is calculated as

$$HR(C) = H(C)/ND$$

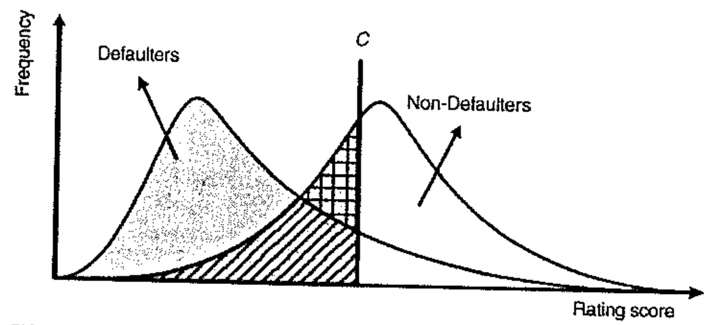


FIGURE 10.1 The Overlap in Internal Ratings Systems
Source: BIS (2005), page 37.

where H(C) is the number of defaulters correctly predicted using cut-off point C, and ND is the number of defaulters in the data sample. The false alarm rate is

$$FAR(C) = F(C)/NND$$

where F(C) is the number of nondefaulters incorrectly forecast as defaulters using cut-off point C (i.e., the cross-hatched area in Figure 10.1) and NND is the number of nondefaulters in the sample.

Figure 10.2 shows how the accuracy of an internal ratings model can be determined by comparing the hit rate to the false alarm rate. The *receiver operating characteristic* (ROC) curve is drawn by plotting the hit rate and false alarm rate for each cut-off point C. The perfect model always has a hit rate of 100 percent and a 0 percent false alarm rate for any cut-off point C. The random model (a 45 degree line) has the accuracy of tossing a coin: 50 percent hit rate and 50 percent false alarm rate. Actual internal ratings models are somewhere between these two extremes, as shown in Figure 10.2. The larger the area under the ROC curve, the more accurate the internal ratings model.[6]

Time-Series versus Cross-Sectional Stress Testing

In a recent set of papers, Granger and Huang (1997), at a theoretical level, and Carey (1998, 2000) and Lopez and Saidenberg (1998), at a simulation/empirical level, show that stress tests similar to those conducted across time for market risk models can be conducted using cross-sectional or panel data for credit risk models. In particular, suppose that in any given year a bank has a sample of N loans in its portfolio, where N is large. By repeated subsampling of the total loan portfolio, it is possible to build up a cross-

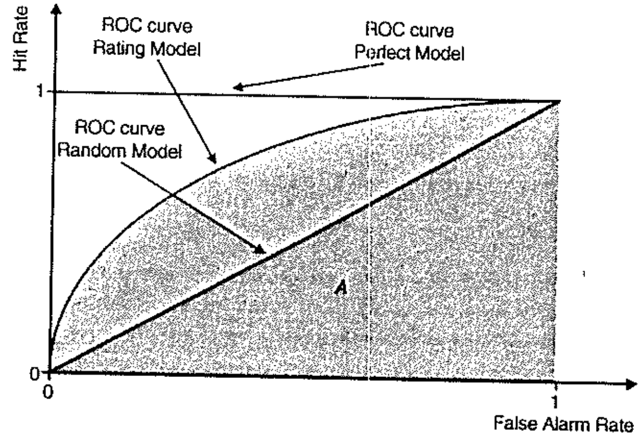


FIGURE 10.2 Measuring the Accuracy of Internal Ratings Models
Source: Bank for International Settlements (2005), page 38.

sectional distribution of expected losses, unexpected losses, and the full probability density function of losses. By comparing cross-sectional sub-portfolio loss distributions with the actual full-portfolio loss distribution, it is possible to generate an idea of the predictive accuracy of a credit risk model. For example, if the model is a good predictor or forecaster, the mean average loss rate and the mean 99th percentile loss rate from 10,000 randomly drawn sub-portfolios of the total loan portfolio should be pretty close to the actual average and 99th percentile loss rates on the full loan portfolio experienced in that year. Indeed, different models may have different prediction errors, and the relative size of the prediction errors can be used to judge the best model (see Lopez and Saidenberg [2000] and Carey [2000]).

A number of statistical issues arise with cross-sectional stress testing, but these are generally similar to those that arise with time-series stress testing (or back-testing). The first issue is that the number of loans in the portfolio has to be large. For example, Carey's (2000) sample is based on 30,000 privately placed bonds held by a dozen life insurance companies during 1986 to 1992, a period during which over 300 credit-related events (defaults, debt restructurings, and so on) occurred for the issuers of the bonds. The subsamples chosen varied in size; for example, portfolios of $0.5 billion to $15 billion in size containing no more than 3 percent of the bonds of any one issuer. Table 10.1 shows simulated loss rates from 50,000 subsample portfolios drawn from the 30,000 bond population. Sub-portfolios were limited to $1 billion in size. Using a Moody's database of

TABLE 10.1 Loss Rate Distribution When Monte Carlo Draws Are From Good versus Bad Years

Portfolio Characteristics		Simulated Portfolio Loss Rates Percent							
			At Loss Distribution Percentiles						
Percent Rated Below BBB	Years Used in Monte Carlo	Mean	95	97.5	99	99.5	99.9	99.95
0%	Good: 1986–1989	0.09	0.53	0.74	1.40	1.46	1.98	2.14
0%	Bad: 1990–1992	0.15	0.87	1.26	1.45	1.59	2.22	2.28
0%	Very bad: 1991	0.16	0.91	1.40	1.54	1.67	2.28	2.36
100%	Good: 1986–1989	1.73	4.18	4.63	5.11	5.43	5.91	6.05
100%	Bad: 1990–1992	2.53	5.59	6.31	7.19	7.82	8.95	9.33
100%	Very bad: 1991	3.76	6.68	7.30	8.04	8.55	9.72	10.19

Source: Carey (1998).

bond ratings and defaults during 1970–1998, Carey (2000) constructs $5 billion subportfolios composed of around 500 bonds and estimates loss distributions under a default mode (DM) model.

The loss rates in Table 10.1 vary by year. In 1991, which was the trough of the last U.S. recession, 50,000 simulated portfolios containing below-investment-grade (rated lower than BBB) bonds produced a (mean) 99 percent loss rate of 8.04 percent, which is quite close to the BIS 8 percent risk-based capital requirement. However, notice that in relatively good years (e.g., 1986–1989), the 99 percent loss rate was much lower at 5.11 percent. Carey (2000) also shows that capital ratios in bad years must be about 175 percent of those in good years if capital is set to cover unexpected losses computed at the 99 percent VAR level.[7]

A related issue is the representativeness of any given year or subperiod chosen to evaluate statistical moments such as the mean (expected) loss rate and the 99 percent unexpected loss rate. Suppose we look at 1991, a recession year. A set of systematic and unsystematic risk factors likely determined the intensity of the recession. The more a recession year reflects systematic rather than unsystematic recession risk factors, the more representative the loss experience of that year is, in a predictive sense, for future bad recession years. This suggests that some type of screening tests need to be conducted on various recession years before a given year's loss experience is chosen as a benchmark for testing predictive accuracy among credit risk models and for calculating capital requirements.[8]

A second issue is the effect of outliers on simulated loss distributions. A few extreme outliers can seriously affect the mean, variance, skewness, and kurtosis of an estimated distribution, as well as the correlations among the loans implied in the portfolio. In a market risk model context, Stahl (1998) has shown how only 5 outliers out of 1,000, in terms of foreign currency exchange rates, can have a major impact on estimated correlations among key currencies. With respect to credit risk, the danger is that a few big defaults in any given year could seriously bias the predictive power of any cross-sectional test of a given model.

Carey (2000) demonstrates the importance of portfolio *granularity* (large disparities in loan sizes within the portfolio) on unexpected loss distributions. Table 10.2 shows that expected losses are relatively unaffected, but that unexpected losses, particularly in the extreme 99.9 percent extreme tails of the distribution, are sensitive to both the size disparity across loans (see rows 1 and 2 of Table 10.2) and large loans to single borrowers (See rows 3 and 4 of Table 10.2).

A third issue deals with variability in LGDs across time and across debt instruments.[9] Table 10.3 shows the wide range of weighted-average LGDs over the period 1978–2001. LGD also varies across industry sectors over time. For example, the telecommunications sector experienced a historically high 88 percent LGD during the second quarter of 2001 (see Altman and Karlin [2001b]). Carey (2000) finds that assumptions about LGD significantly affect the loan portfolio's loss distribution. For example, allowing LGD to vary causes unexpected losses at the 99 percent tail of the loss

TABLE 10.2 The Impact of Loan Size Distribution on Portfolio Losses

Simulation Parameters	Mean	95%	99%	99.5%	99.9%
Base case, 500 loans, random sizes	0.67	2.01	2.98	3.39	4.34
Base case, 500 loans, equal sizes	0.65	1.73	2.37	2.58	2.98
Base case, no one-borrower limit	0.66	2.09	3.38	4.16	7.81
Base case, 5% limit on lending to a single borrower	0.66	2.11	3.14	3.55	4.43

Source: Carey (2001b), Tables 6 and 7.

distribution to increase from 0.64 percent (assuming a fixed LGD of 10 percent for all senior debt and a fixed LGD of 5 percent for all senior debt restructurings) to 3.18 percent for variable LGDs (assuming a mean LGD of 44 percent for senior debt and a mean LGD of 22 percent for senior debt restructurings). Moreover, Fraser (2000) uses CreditMetrics to stress test a portfolio of 331 liquid Eurobonds for LGD sensitivity, finding a significant 0.048 percent increase in portfolio 99 percent VAR for every 1 percent increase in expected LGD.

Stress tests of other model parameters show less sensitivity. For example, Fraser (2000) finds that a 1 percent increase in constant correlations assumed for a Eurobond portfolio causes a 0.026 percent increase in CreditMetrics' estimate of 99 percent VAR, but that the impact was non-monotonic; for certain ranges, as correlations increased, some risk measures actually decreased. Moreover, Carey (2000) finds that the distribution of obligors across industries (with different cross-correlations) does not have much of an impact on unexpected loss estimates.

USING THE ALGORITHMICS MARK-TO-FUTURE MODEL

Back-testing often takes the form of scenario analysis. That is, how will a credit risk model perform under different market scenarios? Stress testing, in particular, focuses on the extreme crisis scenarios. Algorithmics Mark-to-Future (MtF) is a scenario-based model that focuses on estimating each asset's risk and return characteristics under thousands of different scenarios corresponding to all major risk factors, ranging from market risk to operational risk to credit risk. For example, Algorithmics MtF can create 5 to 20 extreme scenarios corresponding to historical market crashes using 50 to 200 systemic market and credit factors in order to conduct credit risk stress tests over time horizons between 1 and 10 years. MtF differs from other credit risk measurement models in that it views market risk and credit risk as inseparable.[10] Stress tests show that credit risk measures are quite sensitive to market risk factors.[11] Indeed, it is the systemic risk parameters that drive creditworthiness in MtF.[12]

Dembo et al. (2000) offer an example of credit risk stress testing using MtF for a BB-rated swap obligation (see Figure 10.3). The firm's credit risk is estimated using a Merton model of default; that is, a creditworthiness index (CWI) is defined that specifies the distance to default as the distance between the value of the firm's assets and a (non-constant) default boundary.[13] Figure 10.3 shows the scenario simulation of the CWI,

TABLE 10.8 Weighted Average (By Issue) Recovery Rates On Defaulted Debt By Seniority Per $100 Face Amount (1978–2009, (3Q))

Default Year	Senior Secured			Senior Unsecured			Senior Subordinated			Subordinated			Discount and Zero Coupon			All Seniorities	
	No.	%	$	No.	%	$	No.	%	$	No.	%	$	No.	%	$	No.	$
2009	27	11%	$38.74	173	73%	$25.90	28	12%	$14.39	4	2%	$12.57	4	2%	$12.23	236	$26.03
2008	18	14%	$30.52	79	63%	$49.56	23	18%	$30.25	4	3%	$21.09	1	1%	$2.71	125	$42.52
2007	10	36%	$87.24	10	36%	$47.70	6	21%	$63.98	2	7%	$46.53	0	0%	$0.00	28	$66.65
2006	9	18%	$90.60	26	52%	$60.90	8	16%	$50.24	1	2%	$60.33	6	12%	$78.31	50	$65.32
2005	67	54%	$76.50	44	36%	$45.88	7	6%	$32.67	0	0%	$0.00	5	4%	$74.21	123	$61.10
2004	27	39%	$63.67	33	48%	$56.77	2	3%	$37.44	0	0%	$0.00	7	10%	$43.06	69	$57.72
2003	57	28%	$53.51	108	53%	$45.40	29	14%	$35.98	1	0%	$38.00	8	4%	$32.27	203	$45.58
2002	37	11%	$52.81	254	75%	$21.82	21	6%	$32.79	0	0%	$0.00	28	8%	$26.47	340	$25.30
2001	9	3%	$40.95	187	67%	$28.84	48	17%	$18.37	0	0%	$0.00	37	13%	$15.05	281	$25.62
2000	13	8%	$39.58	47	29%	$25.40	61	37%	$25.96	26	16%	$26.62	17	10%	$23.61	164	$26.74
1999	14	11%	$26.90	60	47%	$42.54	40	31%	$23.56	2	2%	$13.88	11	9%	$17.30	127	$27.90
1998	6	18%	$70.38	21	62%	$39.57	6	18%	$17.54	0	0%	0.00	1	3%	$17.00	34	$40.46
1997	4	16%	$74.90	12	48%	$70.94	6	24%	$31.89	1	4%	$60.00	2	8%	$19.00	25	$57.61
1996	4	17%	$59.08	4	17%	$50.11	9	38%	$48.99	4	17%	$44.23	3	13%	$11.99	24	$45.44
1995	5	15%	$44.64	9	27%	$50.50	17	52%	$39.01	1	3%	$20.00	1	3%	$17.50	33	$41.77
1994	5	23%	$48.66	8	36%	$51.14	5	23%	$19.81	3	14%	$37.04	1	5%	$5.00	22	$39.44
1993	2	6%	$55.75	7	22%	$33.38	10	31%	$51.50	9	28%	$28.38	4	13%	$31.75	32	$38.83
1992	15	22%	$59.85	8	12%	$35.61	17	25%	$58.20	22	33%	$49.13	5	7%	$19.82	67	$50.03
1991	4	3%	$44.12	69	44%	$55.84	37	24%	$31.91	38	24%	$24.30	9	6%	$27.89	157	$40.67
1990	12	10%	$32.18	31	27%	$29.02	38	33%	$25.01	24	21%	$18.83	11	9%	$15.63	116	$24.66
1989	9	12%	$82.69	16	21%	$53.70	21	28%	$19.60	30	39%	$23.95				76	$35.97
1988	13	21%	$67.96	19	31%	$41.99	10	16%	$30.70	20	32%	$35.27				62	$43.45
1987	4	13%	$90.68	17	55%	$72.02	6	19%	$56.24	4	13%	$35.25				31	$66.63
1986	8	14%	$48.32	11	20%	$37.72	7	13%	$35.20	30	54%	$33.39				56	$36.60
1985	2	7%	$74.25	3	11%	$34.81	7	26%	$36.18	15	56%	$41.45				27	$41.78
1984	4	29%	$53.42	1	7%	$50.50	2	14%	$65.88	7	50%	$44.68				14	$50.62
1983	1	13%	$71.00	3	38%	$67.72				4	50%	$41.79				8	$55.17
1982				16	80%	$39.31				4	20%	$32.91				20	$38.03
1981	1	100%	$72.00													1	$72.00
1980				2	50%	$26.71				2	50%	$16.63				4	$21.67
1979										1	100%	$31.00				1	$31.00
1978				1	100%	$60.00										1	$60.00
Total/ Average	387	15%	$57.55	1279	50%	$36.24	471	18%	$30.04	259	10%	$30.85	161	6%	$25.49	2,557	$36.72
Standard Dev*			$18.39			$13.74			$14.62			$17.46			$20.49		$14.19
Median			$57.42			$45.64			$32.73			$29.69			$18.25		$41.77

*Standard deviations are calculated based on the yearly averages.

Source: Authors' Compilations from Various Dealer Quotes.

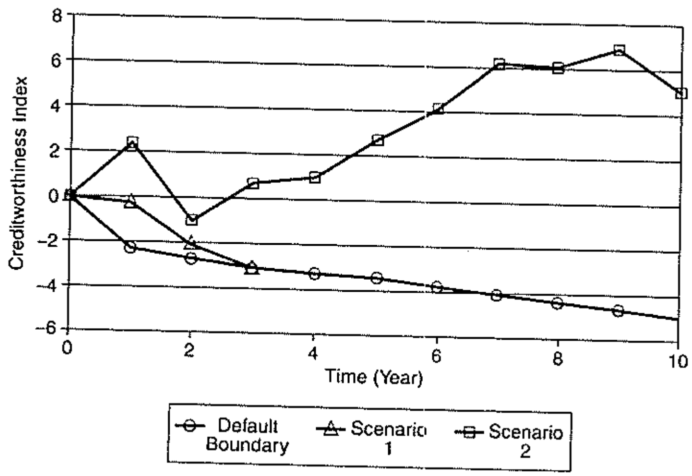


FIGURE 10.3 Merton Model of Default
Source: Dembo et al., (2000), page 68.

illustrating two possible scenarios of firm asset values: In scenario 1 the firm defaults in year 3, while in scenario 2 the firm remains solvent for the next 10 years. The default date under each scenario is represented by the point at which the firm's asset value first hits the default boundary.[14]

MtF assumes that the CWI follows a geometric Brownian motion standardized to have a mean of zero and a variance of one. The basic building block of the CWI is the unconditional cumulative default probability for typical BB-rated firms, obtained using the Merton model (as discussed in Chapter 4). Using the unconditional default probabilities as a foundation, a conditional cumulative default probability distribution is generated for each scenario. That is, the sensitivity of the default probability to scenario risk factors is estimated in the following manner. For example, suppose that the unconditional likelihood of default within five years for a BB firm is 9.6 percent. Choose a particular scenario of the time path of the S&P 500 and six-month U.S. Treasury rates over the next 10 years. This is the credit driver. Suppose that in this particular scenario (call it scenario 9, or S9), the credit driver decreases about 1.2 standard deviations in five years. What is the impact of the decline in the credit driver represented in S9 on the default risk of this BB-rated firm?

MtF estimates all BB-rated firms' historical sensitivity to the credit driver using a multifactor model that incorporates both systemic and

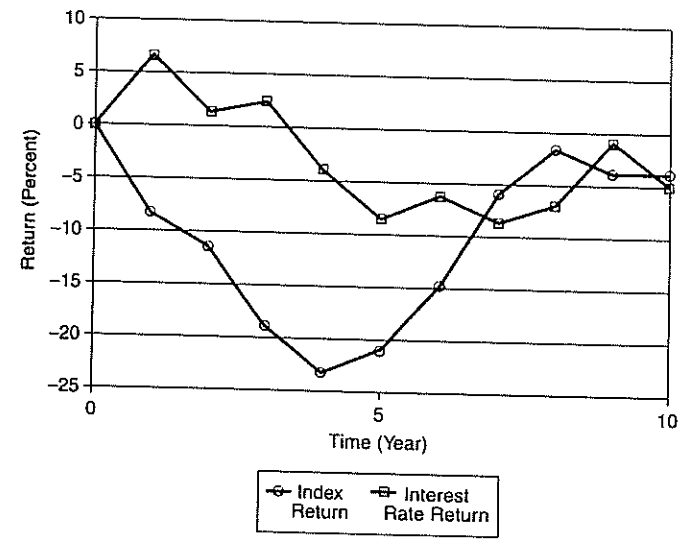


FIGURE 10.4 Scenario S9 Returns
Source: Dembo et al., (2000), page 70.

idiosyncratic credit factors. If the results of the multifactor model suggest that the obligor has a positive correlation to the credit driver, then the swap's credit quality is expected to decrease under this scenario. The conditional cumulative default probability is calculated based on the results of the multifactor model. In this example, the BB firm's five-year probability of default increases from 9.6 percent to 11.4 percent under scenario S9. Figure 10.4 shows the return on the BB swap obligation over the next 10 years using the conditional default probabilities obtained for S9. This process is replicated for several scenarios. Figure 10.5 shows the conditional default probabilities for 10 different credit driver scenarios. A return distribution can be derived using the full range of possible scenarios.

The results for scenario S9 depend on the assumption that systemic risk explains 5 percent of the total variance of the CWI, with idiosyncratic risk explaining the remaining 95 percent. If, by contrast, systemic risk accounted for 80 percent of the variance, the five-year conditional default probability under scenario S9 would have been 44.4 percent instead of 11.4 percent. Therefore, conditional default probabilities have higher volatility when the systemic risk component is greater.

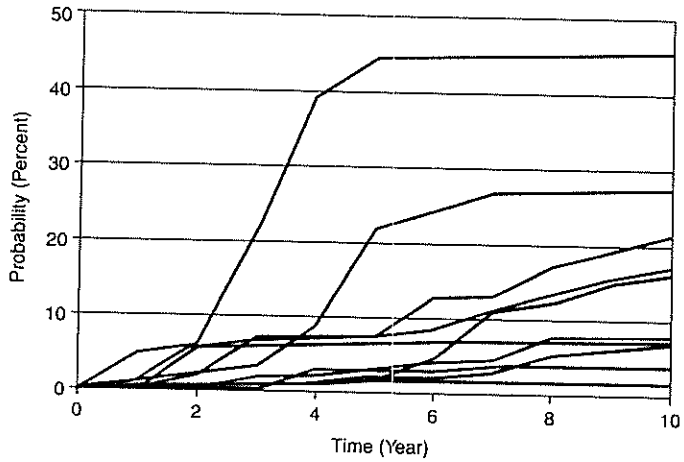


FIGURE 10.5 Ten Scenarios of Conditional Default Probabilities
Source: Dembo et al., (2000), page 70.

STRESS TESTING U.S. BANKS IN 2009

Concerns about the stability of the U.S. banking system in the wake of the financial crisis of 2007–2009 led the Federal Reserve, together with the Federal Deposit Insurance Corporation (FDIC) and the Office of the Comptroller of the Currency (OCC), to require the 19 largest U.S. banks to conduct an unprecedented stress test exercise from February through April 2009 (formally called the Supervisory Capital Assessment Program, or SCAP). All banks with more than $100 billion in year-end assets as of December 2008 were required to participate, representing two-thirds of the assets and more than one half of the loans in the U.S. banking system at the time. Banks were required to conduct a "what if" exercise to forecast their credit losses and revenues under two alternative economic scenarios:

1. The baseline scenario, set equal to the average projections published by Consensus Forecasts, the Blue Chip Survey, and the Survey of Professional Forecasters as of the start of the stress test.
2. The adverse scenario, set by banking supervisors to reflect a longer and deeper recession than expected by market forecasters.

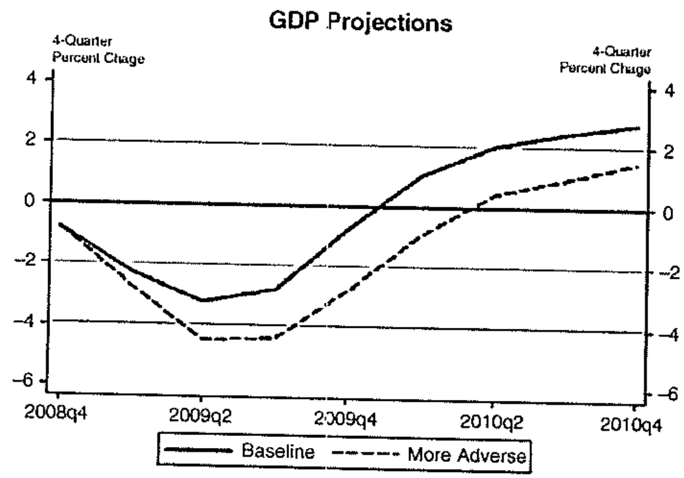
TABLE 10.4 Economic Scenarios: Baseline and More Adverse Alternatives

	2009	2010
Real GDP[a]		
Average Baseline[b]	−2.0	2.1
Consensus Forecasts	−2.1	2.0
Blue Chip	−1.9	2.1
Survey of Professional Forecasters	−2.0	2.2
Alternates More Adverse	−3.3	0.5
Civilian unemployment rate[c]		
Average Baseline[b]	8.4	8.8
Consensus Forecasts	8.4	9.0
Blue Chip	8.3	8.7
Survey of Professional Forecasters	8.4	8.8
Alternative More Adverse	8.9	10.3
House prices[d]		
Baseline	−14	−4
Alternative More Adverse	−22	−7

[a]Percent change in annual average.
[b]Baseline forecasts for real GDP and the unemployment rate equal the average of projections released by Consensus Forecasts, Blue Chip, and Survey of Professional Forecasters in February.
[c]Annual average.
[d]Case-Shiller 10-City Composite, percent change, fourth quarter of the previous year to fourth quarter of the year indicated.
Source: Board of Governors of the Federal Reserve System, *The Supervisory Capital Assessment Program: Design and Implementation*, April 24, 2009, 6.

Table 10.4 shows the macroeconomic assumption in the two alternative scenarios.[15] Figure 10.6 shows the assumed distribution of macroeconomic scenario changes over the two-year forecasting period.

Banks were instructed to calculate their projected losses going forward, not including losses already booked until the end of 2008.[16] They were told to forecast expected losses under the loss reserve provisions of accrual accounting, which require the bank to write down the loan value if repayment becomes doubtful, but not to reflect liquidity-driven declines in market values.[17] There were 12 separate loan categories covered in the stress test: three types of first lien mortgages (prime, Alt-A, and subprime); two types of second/junior lien mortgages (closed-end and home equity lines of



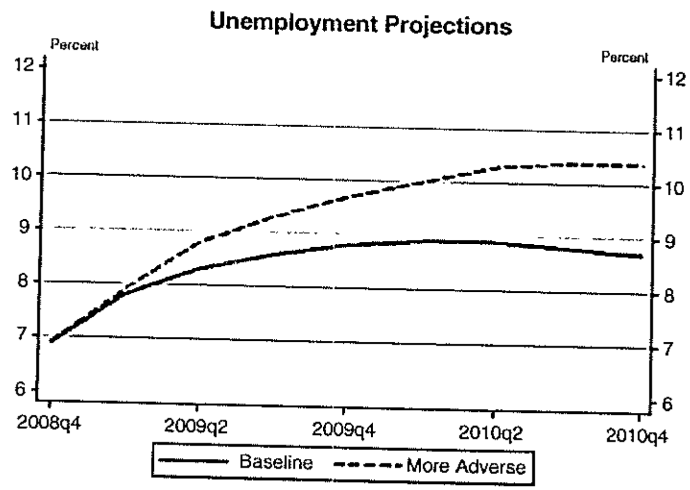


FIGURE 10.6 GDP and Unemployment Projections

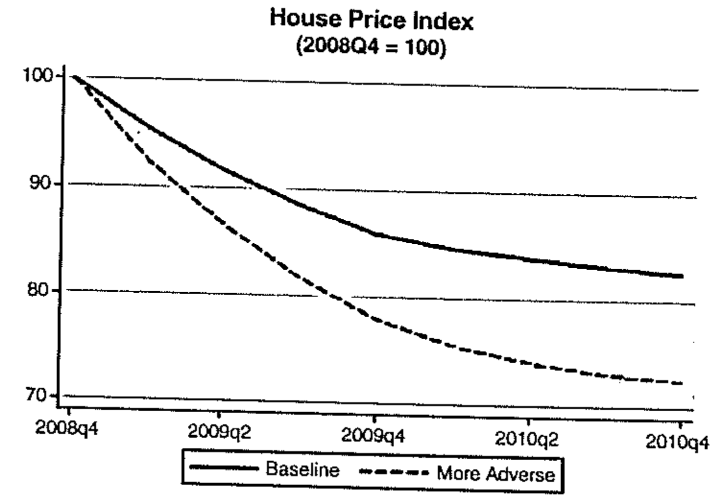


FIGURE 10.6 (*Continued*)
Source: Board of Governors of the Federal Reserve System, *The Supervisory Capital Assessment Program: Design and Implementation*, April 24, 2009, 7.

credit, or HELOCs); commercial and industrial (C&I) loans; three types of commercial real estate (CRE) loans (construction, multifamily, and nonfarm nonresidential); credit cards; other consumer loans; and other loans. For example, the banks were required to provide to regulators detailed data for their residential loan portfolio (separating first mortgages, HELOCs, and closed-end second mortgages) on loan-to-value ratios, FICO scores, geography, documentation level, year of origination, and so on.

Banks were also required to forecast the resources available to cover projected losses. These resources consist of pre-provision net revenue (PPNR) and allowances for loan and lease losses (ALLL) combined with existing capital. The test specified that resources must exceed anticipated losses under both scenarios and still leave sufficient capital to exceed minimum regulatory capital standards. To determine capital requirements, the stress test focused on common stockholders' equity and Tier 1 capital, consisting of common stockholders' equity, qualifying perpetual preferred stock, and certain other assets (subject to limits). The focus of the recommendations was on ascertaining whether the banks had sufficient Tier 1 common stockholders' equity capital to withstand a substantial economic downturn.

TABLE 10.5 Capital Assessment Program for 19 U.S. Bank Holding Companies ($ Billions)

	AmEx	BofA	BB&T	BNYM	CapOne	Citi	FifthThird
Tier 1 capital	10.1	173.2	13.4	15.4	16.8	118.8	11.9
Tier 1 common capital	10.1	74.5	7.8	11.0	12.0	22.9	4.9
Risk-weighted assets	104.4	1,633.8	109.8	115.8	131.8	996.2	112.6
Estimated for 2009 and 2010 for the more adverse scenario							
Total less estimates (before purchase accounting adjustments)	11.2	136.6	8.7	5.4	13.4	104.7	9.1
First lien mortgages	-na-	22.1	1.1	0.2	1.8	15.3	1.1
Second/Junior lien mortgages	-na-	21.4	0.7	-na-	0.7	12.2	1.1
Commercial and industrial loans	-na-	15.7	0.7	0.4	1.5	8.9	2.8
Commercial real estate loans	-na-	9.4	4.5	0.2	1.1	2.7	2.9
Credit card loans	8.5	19.1	0.2	-na-	3.6	19.9	0.4
Securities (AFS and HTM)	-na-	8.5	0.2	4.2	0.4	2.9	0.0
Trading and counterparty	-na-	24.1	-na-	-na-	-na-	22.4	-na-
Other[a] *	2.7	16.4	1.3	0.4	4.3	20.4	0.9
Total Less Rate on Loans[b]	14.3%	10.0%	8.6%	2.6%	11.7%	10.9%	10.5%
First lien morgages	-na-	6.8%	4.5%	5.0%	10.7%	8.0%	10.3%
Second/Junior lien mortgages	-na-	13.5%	8.6%	-na-	19.9%	19.5%	8.7%
Commercial and industrial loans	-na-	7.0%	4.5%	5.0%	9.7%	5.8%	11.0%
Commercial real estate loans	-na-	9.1%	12.6%	9.9%	6.0%	7.4%	13.9%
Credit card loans	20.2%	23.5%	18.2%	-na-	18.2%	23.0%	22.3%
Memo: Purchase accounting adjustments	0.0	13.3	0.0	0.0	1.5	0.0	0.0
Resources other than capital to aboard losses in the more adverse scenario[c]	11.9	74.5	5.5	6.7	9.0	49.0	5.5
SCAP buffer added for more adverse scenario							
(SCAP buffer is defined as additional Tier 1 common/contingent common)							
Indicated SCAP buffer as of December 31, 2008	0.0	46.5	0.0	0.0	0.0	92.6	2.6
Less: Capital actions and effects of Q1 2009 results[d,e,f,g]	0.2	12.7	0.1	−0.2	−0.3	87.1	1.5
SCAP buffer[h,i,j]	0.0	33.9	0.0	0.0	0.0	5.5	1.1

	GMAC	Goldman	JPMC	KeyCorp	MetLife	Morgan Stanley	PNC	Regions	State St	SunTrust	USB	Wells	Total
Tier 1 capital	17.4	55.9	136.2	11.6	30.1	47.2	24.1	12.1	14.1	17.6	24.4	85.4	836.7
Tier 1 common capital	11.1	34.4	87.0	6.0	27.8	17.8	11.7	7.6	10.8	9.4	11.8	33.9	412.5
Risk-weighted assets	172.7	464.8	1,337.5	106.7	326.4	310.6	250.9	116.3	69.6	162.0	230.6	1,082.3	7,814.8
Estimated for 2009 and 2010 for the more adverse scenario													
Total less estimates (before purchase accounting adjustments)	9.2	17.8	97.4	6.7	9.6	19.7	18.8	9.2	8.2	11.8	15.7	86.1	599.2
First lien mortgages	2.0	-na-	18.8	0.1	0.0	-na-	2.4	1.0	-na-	2.2	1.8	32.4	102.3
Second/Junior lien mortgages	1.1	-na-	20.1	0.6	0.0	-na-	4.6	1.1	-na-	3.1	1.7	14.7	83.2
Commercial and industrial loans	1.0	0.0	10.3	1.7	0.0	0.1	3.2	1.2	0.0	1.5	2.3	9.0	60.1
Commercial real estate loans	0.6	-na-	3.7	2.3	0.8	0.6	4.5	4.9	0.3	2.8	3.2	8.4	53.0
Credit card loans	-na-	-na-	21.2	0.0	-na-	-na-	0.4	-na-	-na-	0.1	2.8	6.1	82.4
Securities (AFS and HTM)	0.5	0.1	1.2	0.1	8.3	-na-	1.3	0.2	1.8	0.0	1.3	4.2	35.2
Trading and counterparty	-na-	17.4	16.7	-na-	-na-	18.7	-na-	-na-	-na-	-na-	-na-	-na-	99.3
Other[a] *	4.0	0.3	5.3	1.8	0.5	0.2	2.3	0.8	6.0	2.1	2.8	11.3	83.7
Total Less Rate on Loans[b]	6.6%	0.9%	10.0%	8.5%	2.1%	0.4%	9.0%	9.1%	4.4%	8.3%	7.8%	8.8%	9.1%
First lien morgages	10.2%	-na-	10.2%	3.4%	5.0%	-na-	8.1%	4.1%	-na-	8.2%	5.7%	11.9%	8.8%
Second/Junior lien mortgages	21.2%	-na-	13.9%	6.3%	14.1%	-na-	12.7%	11.9%	-na-	13.7%	8.8%	13.2%	13.8%
Commercial and industrial loans	2.7%	1.2%	6.8%	7.9%	0.0%	2.4%	6.0%	7.0%	22.8%	5.2%	5.4%	4.8%	6.1%
Commercial real estate loans	33.3%	-na-	5.5%	12.5%	2.1%	45.2%	11.2%	13.7%	35.5%	10.6%	10.2%	5.9%	8.5%
Credit card loans	-na-	-na-	22.4%	37.9%	-na-	-na-	22.3%	-na-	-na-	17.4%	20.3%	26.0%	22.5%
Memo: Purchase accounting adjustments	0.0	0.0	19.9	0.0	0.0	0.0	5.9	0.0	0.0	0.0	0.0	23.7	64.3
Resources other than capital to aboard losses in the more adverse scenario[c]	−0.5	18.5	72.4	2.1	5.6	7.1	9.6	3.3	4.3	4.7	13.7	60.0	362.9
SCAP buffer added for more adverse scenario													
(SCAP buffer is defined as additional Tier 1 common/contingent common)													
Indicated SCAP buffer as of December 31, 2008	6.7	0.0	0.0	2.5	0.0	8.3	2.3	2.9	0.0	3.4	0.0	17.3	185.0
Less: Capital actions and effects of Q1 2009 results[d,e,f,g]	−4.8	7.0	2.5	0.6	0.6	6.5	1.7	0.4	0.2	1.3	0.3	3.6	110.4
SCAP buffer[h,i,j]	11.5	0.0	0.0	1.8	0.0	1.8	0.6	2.5	0.0	2.2	0.0	13.7	74.6

[a] Includes other consumer and non-consumer loans and miscellaneous commitments and obligations

[b] Includes losses on other consumer and non-consumer loans

[c] Resources to aboard losses include pre-provision net revenue less the change in the allowance for loan and lease losses

[d] Capital actions include completed or contracted transactions since Q4 2008

[e] For BofA, includes capital benefit from risk-weighted asset impact of eligible asset guarantee

[f] For Citi, includes impact of preferred exchange offers announced on February 27, 2009

[g] Total includes only capital actions and effects of Q1 2009 results for firms that need to establish a SCAP buffer

[h] There may be a need to establish an additional Tier 1 capital buffer, but this would be satisfied by the additional Tier 2 Common capital buffer unless otherwise specified for a particular BHC

[i] GMAC needs to augment the capital buffer with $11.5 billion of Tier 1 Common/contingent Common of which $9.1 billion must be new Tier 1 capital

[j] Regions needs to augment the capital buffer with $2.5 billion of Tier 2 Common/contingent Common of which $400 million must be new Tier 2 capital

Note: Numbers may not sum up to 1 due to rounding.

Source: Board of Governors of the Federal Reserve System, *The Supervisory Capital Assessment Program: Design and Implementation*, April 24, 2009, 9.

The results of the stress test show that the aggregate losses at the top 19 U.S. banks could equal $600 billion during 2009 and 2010 under the adverse economic scenario. Although the aggregate resources available to meet these losses was estimated at $835 billion, additional capital totaling $74.6 billion was required for 10 of the 19 banks. Therefore, 9 of the banks were found to have sufficient capital to withstand the adverse economic scenario, and the banking system as a whole (as measured by the largest 19 banks) was found to be fundamentally solvent.

Table 10.5 shows the results of the stress test for each of the 19 banks. The banks that passed the stress test (i.e., those which required no additional SCAP buffer as shown in the last row of Table 10.5) were American Express, BB&T Corporation, Bank of New York Mellon, Capital One Financial Corporation, Goldman Sachs Group, JPMorgan Chase, MetLife, State Street Corporation and U.S. Bancorp. Figure 10.7 shows each bank's total projected losses as a fraction of year-end 2008 risk-weighted assets.[18]

Banks with capital deficiencies, according to the stress test scenario, were required to provide a plan for resolving their deficiencies within 30 days, to be implemented within six months. The banks that passed the stress test quickly (as of June 2009) repaid the funds granted to them in October 2008 under the Troubled Asset Relief Program (TARP), see Chapter 3. However, companies such as JPMorgan Chase and Morgan Stanley have to

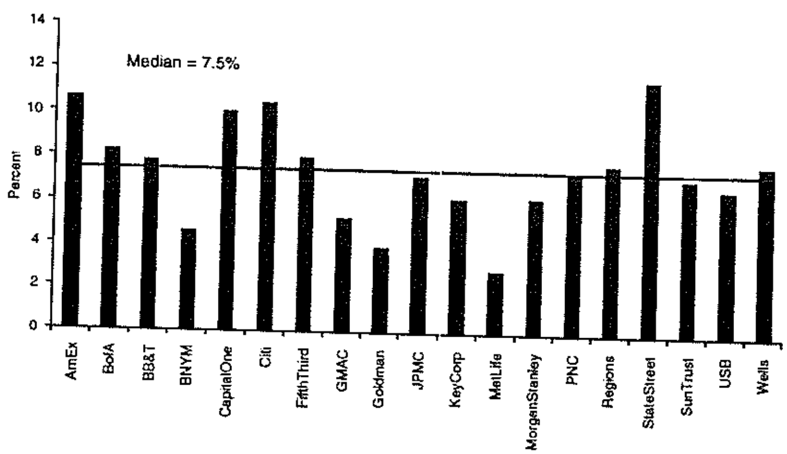


FIGURE 10.7 Supervisor Estimates of Total Losses to Risk-Weighted Assets for More Adverse Scenario

Source: Board of Governors of the Federal Reserve System, *The Supervisory Capital Assessment Program: Overview of Results*, May 7, 2009, 10.

find a price to repurchase the warrants issued to the government in order to completely exit the TARP.

It might be noted that this stress test exercise has not been without its critics. In particular, by defining capital measures that include perpetual preferred stock and various intangibles and "other assets," the capital measures were in all probability biased upwards. Moreover, the risk-weighted assets denominator had credit risk weights reflecting the precrisis exposures of securitized assets as per the 2006 Basel II model. This would have biased the denominator of the Tier 1 capital ratio downwards and the calculated capital ratios upwards.

SUMMARY

A key measure of the usefulness of internal credit risk models is their predictive ability. Tests of predictive ability, such as back-testing, are difficult for credit risk models because of the lack of data for a sufficiently long time series. Nevertheless, given a large and representative (in a default risk sense) loan portfolio, it is possible to stress test credit risk models by using cross-sectional subportfolio sampling techniques that provide predictive information on average loss rates and unexpected loss rates. Moreover, the predictive accuracy, in a cross-sectional sense, of different models can be used to choose among different models. In the future, wider-panel data sets and even time-series of loan loss experience are likely to be developed by banks and/or consortia of banks.

Another approach to credit risk stress testing that avoids the problem of data limitations is the scenario analysis approach, such as that adopted by Algorithmics Mark-to-Future. Credit drivers, composed of market risk factors, are used to estimate conditional default probabilities. Varying the credit driver scenario causes changes in conditional default probabilities which are then used to determine a creditworthiness index. Scenarios can also be chosen to replicate extreme events in order to stress test the portfolio's credit risk exposure.

The stress tests conducted by bank regulators during the early part of 2009 consisted of a forward-looking forecast of credit losses and revenues under several economic scenarios. The results showed that the aggregate worst-case losses under adverse economic conditions could be met by resources (revenues and shareholders' equity) for the largest 19 companies in the U.S. banking system. However, 9 out of 19 banks were required to raise additional capital totaling $75 billion.