Confidence intervals for corporate default rates

Rating agency default studies provide estimates of mean default rates over multiple time horizons but have never included estimates of the standard errors of the estimates. This is due at least in part to the challenge of accounting for the high degree of correlation induced by their cohort-based methodologies. In this article, **Richard Cantor, David Hamilton** *and* **Jennifer Tennant** *present a method for estimating confidence intervals for corporate default rates derived through a bootstrapping approach*

Historical average cumulative default rates by rating category and investment horizon are among rating agencies' most widely referenced statistics. In any finite sample, however, the historical mean default rate may overstate or understate the underlying population's true risk of default, depending upon whether the particular set of issuers included in the sample happen to experience lower or higher than expected default incidence. Quantitative credit analysts and risk managers are, therefore, not only interested in mean default rates but also in their standard errors, which they can use to calculate confidence intervals around the estimated means.

Banks that map their internal rating systems to agency rating scales are also interested in the confidence intervals around the average default rates by rating category. In particular, banks may want to know whether the default rates associated with the so-called 'low-default portfolio' portion of the rating scale (A, Aa and Aaa credits) are sufficiently differentiated by rating category to meet Basel II criteria for internal ratings systems.

In this article, we estimate confidence intervals for Moody's average default rates by rating category and investment horizon based on corporate issuer rating histories from 1970–2006 using a non-parametric bootstrapping approach. The work extends research in the academic literature on one-year default rates (Hanson & Schuermann, 2006) to the multi-year horizon case. We find that, for the estimated mean default rates associated with the Baa and speculative-grade rating categories, the standard errors are generally small, at 10% or less than the mean. The mean default rates for the Aaa, Aa and A rating categories, however, tend to be less precisely estimated, with standard errors ranging from 15% to about 100% of the means, depending on the rating category and the investment horizon. We also find that the precision of the estimate of the mean tends to increase with the length of the investment horizon, because the mean default rate rises faster than the standard error as the horizon increases.

The results suggest that Moody's long-term ratings satisfy the Basel II criteria for effectively distinguishing relative credit risk. At the one-year horizon, default rates by rating category are properly rank ordered, but in some cases their differences may not be significantly different in the statistical sense. At longer horizons, however, their differences are clearly significantly different. This is true even for the Aaa, Aa and A rating categories, as the default rates associated with these rating categories are significantly different from one another at the two-year (and longer) horizon.[1]

As a byproduct of this analysis, we also present confidence intervals for one-year hazard rates, which in rating agency parlance are known as marginal default rates. The marginal default rate of year t is the probability of default in the tth year conditional on not defaulting (or having a rating withdrawn) prior to year t. Interestingly, as t increases and the horizon lengthens, the precision of the estimated mean marginal default rates (as measured by the inverse of the coefficient of variation) improves more slowly than does the cumulative default rate, and even decreases in many cases.

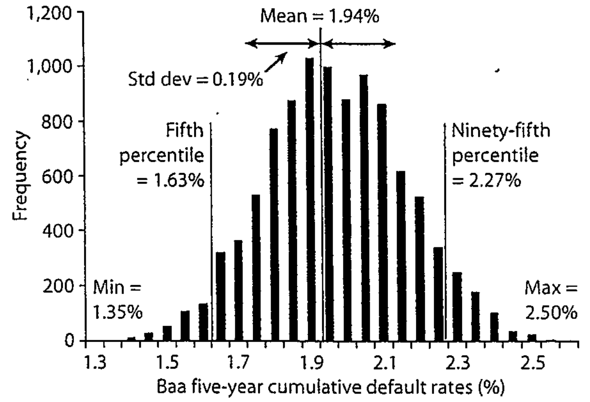
It is important to emphasise, however, that standard errors around estimated long-run average default rates should not be confused with the much greater bands of uncertainty associated with expected performance of particular cohorts of issuers formed at specific points in time. Long-term default rates average across the performance of many macroeconomic scenarios and credit market environments, whereas the default rate associated with any particular cohort will depend upon the realisation of a particular macroeconomic scenario and credit market environment.

Methods for measuring long-run mean default rates and their confidence intervals

The academic literature suggests multiple approaches for estimating confidence intervals around historical average one-year default rates: analytic, parametric and non-parametric (that is, bootstrap)

[1] These long-term horizon differences can be used to validate the effectiveness of these rating categories in distinguishing risk for internal rating systems. Hence, they can be used to differentiate risk within internal rating systems (see Basel Committee on Banking Supervision, 2005, page 5)

1 Frequency distribution of Baa five-year cumulative default rates in 10,000 bootstrapped samples



methods.[2] The simplest possible analytic approach assumes that default events are independent, and whether or not an issuer defaults is determined by a binomial process governed by a single known parameter (the probability of default). In large samples, the probability distribution of the empirical average default rate is known to be normally distributed with a mean equal to the population's underlying default rate and a standard deviation that is a decreasing function of the size of the sample.[3]

Another approach is to specify and estimate parameters governing the evolution of the default process over time. By adding structure, such parametric methods use the available data very efficiently; however, they can also be misleading if they are based on inaccurate models of the default process. In fact, in common applications, most parametric models appear to be based on assumptions directly at odds with reality and therefore probably yield misleading or invalid estimates.[4]

As a practical matter, even if parametric and analytic approaches were to prove useful for the analysis of short-term default rates, they would remain difficult to adapt to the analysis of multi-year cumulative default rates. To calculate long-term default rates from a mix of short-term and long-term rating histories, Moody's (but not all other rating agencies) follows the standard academic practice of adjusting its default rates for the censoring introduced by short-lived rating histories, which if left unadjusted lead to artificially low estimates of long-term default rates. As discussed in Cantor & Hamilton (2007), Moody's calculates long-term default rates by compounding one-year marginal default rates derived from cohort data constructed at monthly intervals.[5] This discrete-time method maximises the information available in the data set; however, it implies that the sample size for a multi-year default rate estimate is not easily defined, since the multi-year default rate is derived from a sequence of one-year marginal default rates, each with its own sample size.

In contrast, a non-parametric bootstrap method makes no assumptions about the statistical process generating the data, except that each issuer's rating and default history are independent.[6] Standard error estimates and other statistical inferences can be made based on Monte Carlo sampling using the actual distribution of the historical data. Bootstrap methods are equally appropriate for estimating confidence intervals for one-year and multi-year cumulative default rates.[7]

Bootstrap estimates of multi-period default rates and their confidence intervals

Our bootstrap analysis follows closely the approach taken in Hanson & Schuermann (2006), except we extend their analysis of one-year default rates to include multi-year cumulative default rates.[8] We create 10,000 data sets by sampling with replacement from our original historical data set of 11,370 firm rating and default histories. In essence, we generate 10,000 cohort default studies, so the results reflect not just one realisation of history but 10,000; each of the 10,000 data sets itself has 11,370 firm rating and default histories.[9] As a result, a firm that naturally appears only once in the original data set will not appear at all in some data sets and will appear multiple times in other data sets – so that across all data sets each firm will on average appear 10,000 times.

The bootstrapped data sets can be used to construct 10,000 realisations on any particular sample statistic. For example, figure 1 depicts the frequency distribution of all 10,000 of the estimated five-year cumulative default rates for issuers rated Baa on each cohort's formation date. Notice that the mean for this sample is 1.94%, which is the same as that reported for the five-year Baa default rate in Moody's annual default study (Hamilton, 2007). The distribution around the mean, however, reveals that resampling 10,000 times from the original data set could conceivably have produced a default rate as low as 1.35% and as high as 2.50%. Yet both of these outcomes are extraordinarily unlikely. The 5% and 95% confidence level points are more meaningful measures of the range of potential outcomes we are ever likely to see. As revealed in figure 1, 95% of the time we would expect to see a long-term average five-year Baa default rate at least equal to 1.63% and no greater than 2.27%. Figure 1 also indicates the standard deviation, 0.19%, of estimated default rate.

Table A presents sample statistics for Moody's letter rating categories (Aaa through Caa) over time horizons ranging from one to 10 years. The table indicates that most rating categories at most time horizons have significantly different estimated mean default rates. Some of the best rating categories (Aaa, Aa and A – the low-default portfolio ratings in Basel terminology) are not statistically distinguishable at the one-year horizon. They are, however, clearly distinguishable at the two-year horizon, as the confidence interval from the fifth to the ninety-fifth percentile ranges from 0.000% to 0.000% for Aaa, from 0.003%

[2] *For a summary and analysis of methods for estimating confidence intervals for the long-term average one-year default rate, see Hanson & Schuermann (2006)*

[3] *The utility of this statistic in measuring one-year default rate confidence intervals is discussed in Cantor & Falkenstein (2001), Stein (2006), and Hanson & Schuermann (2006). Cantor & Falkenstein (2001) also discuss how the approach can be adapted to take into account common temporal shocks to the aggregate default rate, which is equivalent to the use of the correlated binomial distribution (see Witt, 2004)*

[4] *For example, Hanson & Schuermann (2006) show that parametric continuous time estimators imply much lower default rates than observed in practice. This result is unsurprising, since commonly used continuous time estimators assume rating transitions are Markovian, an assumption that is inconsistent with the rating change momentum that is strongly evident in the data (see Hamilton & Cantor, 2004). The advantages and practical limitations of deriving non-Markovian continuous time estimators are explored by Christensen, Hansen & Lando (2004)*

[5] *A cohort is simply a portfolio of issuers holding the same rating on a given date, for example, the Baa-rated cohort formed on January 1, 2006. Each cohort is tracked from the time of formation into the future, and each default and rating withdrawal is recorded. By forming cohorts at periodic (in our case, monthly) intervals, one can calculate expected default rates. The cohort method is a discrete-time approximation of the continuous time hazard rate method*

[6] *The assumption of the independence of issuer rating histories is, strictly speaking, violated in reality. However, the assumption minimises any correlation in default times that may exist in the data. For example, the default time for an issuer first rated in, say, 1974 is unlikely to be strongly correlated with the default time for an issuer whose rating history begins in 1998*

[7] *Our intention in this article is not to compare these methods, but to demonstrate that confidence intervals for long-run default rates exhibit sufficiently differentiated performance using a very straightforward statistical method*

[8] *Moreover, our data sample is almost twice as large; therefore, our confidence intervals for the one-year default rates are considerably tighter than their bootstrap results*

[9] *The choice of 10,000 iterations was sufficient to begin to generate statistically significant differences in default rates among rating categories at time horizons longer than one year*

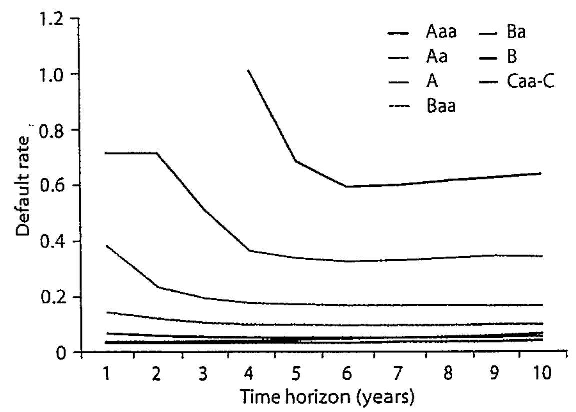
A: Historical average cumulative default rates and bootstrapped confidence intervals (bp)

		Time horizon (years)									
		1	2	3	4	5	6	7	8	9	10
Aaa	Mean (μ)	0.00	0.00	0.00	2.72	10.19	17.51	25.36	33.76	42.75	52.43
	Standard deviation (σ)	0.00	0.00	0.00	2.75	6.97	10.40	15.20	20.74	26.84	33.49
	Precision (σ/μ)	–	–	–	1.01	0.68	0.59	0.60	0.61	0.63	0.64
	Fifth percentile	0.00	0.00	0.00	0.00	1.00	4.22	5.09	5.09	5.09	5.09
	Ninety-fifth percentile	0.00	0.00	0.00	7.99	23.26	36.21	52.94	71.47	91.10	113.55
	Minimum	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Maximum	0.00	0.00	0.00	17.97	49.89	74.04	99.93	127.72	158.73	197.23
Aa	Mean (μ)	0.78	1.87	4.19	10.55	17.67	25.88	34.17	41.42	46.20	52.04
	Standard deviation (σ)	0.56	1.34	2.16	3.84	5.94	8.44	11.23	13.92	15.87	17.84
	Precision (σ/μ)	0.72	0.72	0.51	0.36	0.34	0.33	0.33	0.34	0.34	0.34
	Fifth percentile	0.00	0.26	1.03	4.76	8.52	12.85	16.85	19.87	21.79	24.61
	Ninety-fifth percentile	1.79	4.42	8.19	17.27	27.88	40.98	54.58	66.63	74.54	83.62
	Minimum	0.00	0.00	0.00	0.79	1.96	4.25	7.19	7.55	7.55	7.55
	Maximum	3.35	7.57	13.35	30.95	47.25	61.78	79.05	99.58	115.22	130.40
A	Mean (μ)	2.08	9.53	22.11	34.47	47.29	61.52	76.06	92.67	110.69	128.79
	Standard deviation (σ)	0.79	2.25	4.30	6.10	8.13	10.49	12.92	15.55	18.44	21.40
	Precision (σ/μ)	0.38	0.24	0.19	0.18	0.17	0.17	0.17	0.17	0.17	0.17
	Fifth percentile	0.91	6.09	15.34	24.62	34.01	44.75	59.91	67.62	81.87	95.49
	Ninety-fifth percentile	3.45	13.42	29.26	44.87	61.26	79.40	98.22	119.48	141.81	164.79
	Minimum	0.07	2.62	8.60	15.90	23.72	32.51	38.32	46.79	54.73	63.83
	Maximum	4.94	17.53	37.93	58.81	79.04	101.07	123.00	147.66	172.62	205.99
Baa	Mean (μ)	18.16	50.80	93.20	143.73	194.36	245.72	296.54	345.91	402.62	464.85
	Standard deviation (σ)	2.58	6.24	10.14	14.55	19.08	23.72	28.65	33.80	39.58	46.18
	Precision (σ/μ)	0.14	0.12	0.11	0.10	0.10	0.10	0.10	0.10	0.10	0.10
	Fifth percentile	14.02	40.82	76.94	119.77	163.07	206.88	249.67	290.40	338.66	391.10
	Ninety-fifth percentile	22.54	61.43	110.09	168.24	226.81	285.36	344.02	400.52	467.20	541.33
	Minimum	10.09	31.26	59.69	95.49	135.08	176.45	206.85	236.40	279.52	327.37
	Maximum	28.10	70.27	125.49	187.86	250.28	315.78	388.99	461.27	537.73	623.63
Ba	Mean (μ)	120.20	321.84	556.55	795.30	1,021.09	1,223.59	1,400.79	1,571.31	1,739.88	1,912.68
	Standard deviation (σ)	7.98	18.56	30.15	41.73	53.14	63.53	73.13	83.13	93.46	104.36
	Precision (σ/μ)	0.07	0.06	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
	Fifth percentile	107.31	291.18	508.07	726.28	934.06	1,120.27	1,284.84	1,436.86	1,588.75	1,744.06
	Ninety-fifth percentile	133.22	352.05	605.15	863.71	1,108.66	1,329.86	1,521.78	1,707.79	1,893.41	2,084.77
	Minimum	88.58	255.01	455.02	660.56	843.04	1,010.45	1,136.76	1,260.07	1,381.05	1,495.11
	Maximum	155.38	400.51	681.40	965.20	1,233.60	1,464.60	1,679.50	1,881.26	2,072.16	2,264.35
B	Mean (μ)	523.95	1,131.19	1,706.61	2,208.26	2,682.24	3,100.41	3,478.49	3,798.86	4,092.44	4,333.87
	Standard deviation (σ)	18.38	37.56	55.64	72.69	90.29	107.02	124.00	140.86	160.10	179.12
	Precision (σ/μ)	0.04	0.03	0.03	0.03	0.03	0.03	0.04	0.04	0.04	0.04
	Fifth percentile	493.86	1,068.52	1,614.88	2,089.37	2,534.16	2,923.33	3,275.51	3,568.14	3,837.63	4,046.91
	Ninety-fifth percentile	554.63	1,193.97	1,798.79	2,327.05	2,830.35	3,279.53	3,681.01	4,034.77	4,358.64	4,633.32
	Minimum	462.64	1,017.38	1,549.17	1,988.85	2,407.52	2,762.69	3,070.43	3,356.00	3,571.55	3,720.78
	Maximum	589.94	1,261.43	1,896.82	2,453.36	2,982.26	3,446.40	3,863.16	4,251.90	4,629.74	4,993.19
Caa-C	Mean (μ)	1,948.83	3,052.50	3,973.68	4,694.87	5,268.13	5,687.13	6,000.10	6,332.61	6,638.68	6,934.08
	Standard deviation (σ)	74.70	117.89	159.00	195.87	233.29	269.09	304.39	350.11	384.45	461.57
	Precision (σ/μ)	0.04	0.04	0.04	0.04	0.04	0.05	0.05	0.06	0.06	0.07
	Fifth percentile	1,827.04	2,860.26	3,707.43	4,370.85	4,800.52	5,142.09	5,353.73	–	–	–
	Ninety-fifth percentile	2,071.19	3,259.05	4,253.14	5,036.36	5,753.29	6,230.03	6,625.56	–	–	–
	Minimum	1,684.18	2,626.62	3,459.97	4,067.36	4,523.00	4,801.66	5,062.66	5,195.15	5,310.82	5,310.82
	Maximum	2,291.66	3,480.96	4,522.24	5,417.92	6,147.27	6,625.15	7,004.56	7,455.23	8,009.47	8,636.60

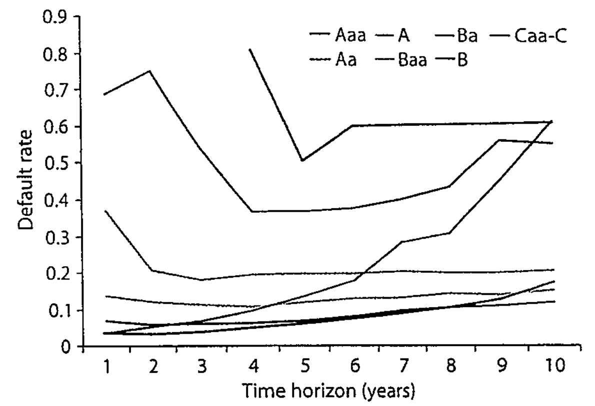
B Historical average marginal default rates and bootstrapped confidence intervals (bp)

		Time horizon (years)									
		1	2	3	4	5	6	7	8	9	10
Aaa	Mean (μ)	0.00	0.00	0.00	2.72	7.47	7.33	7.87	8.44	9.04	9.73
	Standard deviation (σ)	0.00	0.00	0.00	2.19	3.77	4.40	4.74	5.11	5.50	5.95
	Precision (σ/μ)	–	–	–	0.81	0.51	0.60	0.60	0.61	0.61	0.61
	Fifth percentile	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00
	Ninety-fifth percentile	0.00	0.00	0.00	7.99	16.13	16.75	18.02	19.33	20.82	22.50
	Minimum	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Maximum	0.00	0.00	0.00	17.97	35.52	29.94	31.98	34.12	36.43	39.12
Aa	Mean (μ)	0.78	1.09	2.33	6.36	7.13	8.23	8.31	7.28	4.80	5.87
	Standard deviation (σ)	0.53	0.82	1.26	2.34	2.62	3.08	3.32	3.15	2.69	3.26
	Precision (σ/μ)	0.69	0.75	0.54	0.37	0.37	0.37	0.40	0.43	0.56	0.55
	Fifth percentile	0.00	0.00	0.31	2.75	3.07	3.44	2.80	2.26	1.15	0.00
	Ninety-fifth percentile	1.79	2.73	4.74	10.48	11.95	13.80	14.33	12.90	9.67	11.82
	Minimum	0.00	0.00	0.00	0.00	0.00	0.91	0.00	0.00	0.00	0.00
	Maximum	3.35	4.71	8.67	19.37	18.15	21.49	21.73	20.83	16.94	23.22
A	Mean (μ)	2.08	7.45	12.58	12.40	12.86	14.30	14.63	16.74	18.19	18.31
	Standard deviation (σ)	0.77	1.57	2.28	2.41	2.53	2.81	2.98	3.36	3.65	3.80
	Precision (σ/μ)	0.37	0.21	0.18	0.19	0.20	0.20	0.20	0.20	0.20	0.21
	Fifth percentile	0.91	4.83	8.95	8.66	8.97	9.84	9.97	11.20	12.19	11.78
	Ninety-fifth percentile	3.45	10.32	16.40	16.37	17.17	19.23	19.71	22.65	24.75	25.01
	Minimum	0.07	2.12	5.66	5.34	5.62	4.69	5.69	6.94	7.46	7.20
	Maximum	4.94	13.71	21.94	16.37	21.89	24.86	26.18	31.39	34.31	35.24
Baa	Mean (μ)	18.16	32.70	42.62	51.01	51.38	52.38	52.11	50.90	58.76	64.87
	Standard deviation (σ)	2.52	3.94	4.78	5.47	6.09	6.76	6.90	7.24	8.28	9.81
	Precision (σ/μ)	0.14	0.12	0.11	0.11	0.12	0.13	0.13	0.14	0.14	0.15
	Fifth percentile	14.02	26.27	35.04	41.83	41.65	42.19	41.08	39.35	45.40	49.62
	Ninety-fifth percentile	22.54	39.49	50.33	60.16	61.45	62.75	63.77	63.29	73.23	80.45
	Minimum	10.09	19.37	28.05	34.36	33.15	32.07	30.95	29.62	29.11	34.82
	Maximum	28.10	46.18	57.61	70.99	71.60	77.15	79.69	80.21	90.93	107.08
Ba	Mean (μ)	120.20	204.10	242.53	252.86	245.35	225.60	201.99	198.40	200.13	209.35
	Standard deviation (σ)	8.37	12.10	14.85	16.29	17.21	18.32	19.64	20.65	22.27	24.77
	Precision (σ/μ)	0.07	0.06	0.06	0.06	0.07	0.08	0.10	0.10	0.11	0.12
	Fifth percentile	107.31	184.30	218.64	226.01	217.80	197.75	172.48	166.38	164.12	169.29
	Ninety-fifth percentile	133.22	224.22	265.97	279.92	275.09	255.28	232.56	231.18	235.32	249.66
	Minimum	88.58	162.02	190.11	203.32	185.99	160.44	134.10	135.36	124.54	132.33
	Maximum	155.38	256.59	300.42	312.58	308.94	289.78	259.17	268.38	283.01	302.42
B	Mean (μ)	523.95	640.85	648.90	605.03	608.54	571.78	548.43	491.85	474.26	409.73
	Standard deviation (σ)	18.68	22.63	25.96	30.40	36.94	42.36	49.10	51.47	59.94	70.60
	Precision (σ/μ)	0.04	0.04	0.04	0.05	0.06	0.07	0.09	0.10	0.13	0.17
	Fifth percentile	493.86	602.11	604.91	556.57	551.18	507.70	474.04	411.58	375.30	302.77
	Ninety-fifth percentile	554.63	679.87	694.03	656.04	668.07	638.88	629.33	579.72	580.41	524.07
	Minimum	462.64	567.14	569.21	507.13	500.58	436.91	382.45	320.18	265.52	204.55
	Maximum	589.94	726.08	746.15	716.43	730.06	722.87	701.67	672.34	786.24	676.97
Caa-C	Mean (μ)	1,948.83	1,371.33	1,327.12	1,198.83	1,083.97	890.42	732.10	842.59	844.58	906.92
	Standard deviation (σ)	72.38	73.63	93.02	117.41	145.55	160.19	206.28	257.78	382.48	555.52
	Precision (σ/μ)	0.04	0.05	0.07	0.10	0.13	0.18	0.28	0.31	0.45	0.61
	Fifth percentile	1,827.04	1,241.84	1,159.17	992.34	782.78	545.23	335.09	–	–	–
	Ninety-fifth percentile	2,071.19	1,512.22	1,502.09	1,414.65	1,417.40	1,281.25	1,205.24	–	–	–
	Minimum	1,684.18	1,133.31	1,047.59	843.57	639.64	382.91	129.74	97.26	0.00	0.00
	Maximum	2,291.66	1,653.44	1,739.77	1,679.39	1,683.53	1,593.59	1,616.99	2,055.70	2,946.91	4,580.15

2 Ratio of the standard deviation to the mean cumulative default rates for the bootstrapped sample



3 Ratio of the standard deviation to the mean marginal default rates for the bootstrapped sample



for 0.044% for Aa and from 0.061% to 0.134% for A. All Moody's long-term ratings would therefore meet the Basel II back-testing criteria for internal rating systems (Basel Committee on Banking Supervision, 2005, page 5).

Figure 2 plots the ratios of the standard deviations to the means for these cumulative default rates (the coefficients of variation), which are inversely related to the precision with which these means are estimated. As seen in figure 2, for the Baa and the speculative-grade rating categories, the standard error is generally small, at 10% or less than the mean. The mean default rates for the Aaa, Aa and A rating categories, however, tend to be less precisely estimated, with standard errors ranging from 15% to about 100% of their respective means. Figure 2 shows that means are more precisely estimated for the higher risk rating categories and at longer time horizons. These results are consistent with the well-known finding that accurate default rate estimation requires larger data sets when the default rates are lower, that is, at better rating categories and at shorter horizons.

It is worth reiterating that Moody's cumulative default rates are adjusted for data censoring, that is, defaults and rating withdrawals that occur prior to the current time interval.[10] Hence, the effective sample size (and the size of the denominators of the marginal default rates) falls as the time horizon lengthens. The decrease in the effective sample size due to adjustments for data censoring is often misunderstood as inducing an upward bias in average cumulative default rate estimates.[11] However, this view confuses concerns about sample size and statistical significance with bias.

Interestingly, the precision of the default rate estimates does not deteriorate at long time horizons, such as 10 years, even though long-term default rates are cumulated from one-year marginal default rates at each horizon, and in some of the cohorts, the number of observations in the tenth year are quite small.[12] The small samples in some cohorts' later years do not cause a problem for the overall sample estimates, because the overall estimates are derived from averages across a large number of cohorts.[13] Table A and figure 2 reveal that while the standard deviations of default rates by rating category do indeed increase as time horizon lengthens, the precision of the estimates of the mean default rates actually increases with the length of the time horizon.

Bootstrap estimates of marginal default rates and their confidence intervals

Cumulative default rates are, of course, measured more precisely than marginal default rates. That is, the estimate of the probability of default over five years is unsurprisingly more precisely estimated than the probability of default in the fifth year, conditional on having survived until the start of the fifth year.

Table B and figure 3 present the same sample statistics for marginal default rates and plots their coefficients of variation (the inverse of the precision of the mean estimate). Interestingly, as the horizon lengthens, the precision of the estimated mean marginal default rates (as measured by the inverse of the coefficient of variation) improves more slowly than does the precision of the cumulative default rate and, in fact, decreases in many cases.

Uncertainty around long-run average CDRs for individual cohorts

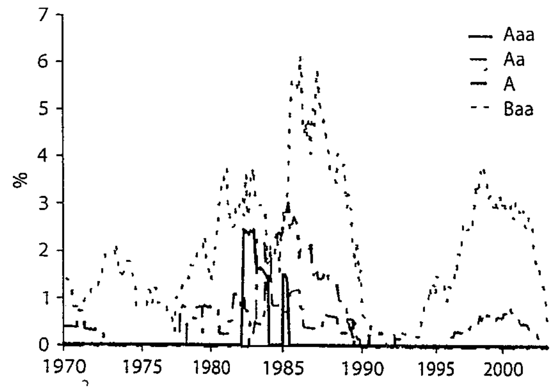
In the previous section, we presented estimates of uncertainty about the long-run average mean. It should be obvious, however, that these measures grossly understate uncertainty about the performance of any particular rating cohort. The long-run average default rates smooth out variations in macroeconomic conditions, whereas the measured performance of any one cohort will reflect the unique macroeconomic conditions that prevailed during its 'lifetime'.

The high variability of cohort default rates should not be surprising because the primary intent of agency ratings (Moody's ratings in particular) is to achieve a powerful rank ordering of credit risk within the current cohort, and the secondary intent is to maintain rating stability when possible.[14] Rating agencies do not shift all their ratings up or down with changes in macroeconomic conditions, and, hence, wide variations in default rates by rating category should be expected over time.

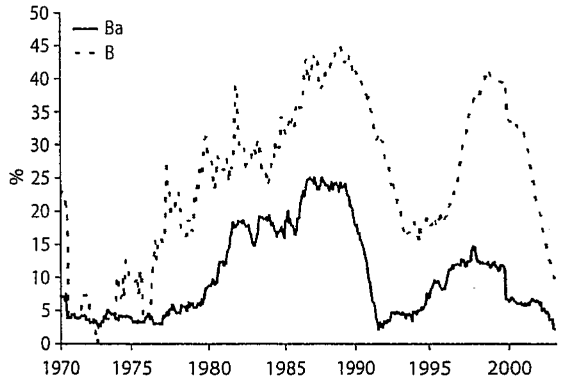
Figures 4 and 5 present the historical realisations of five-year default rates for investment-grade and speculative-grade cohorts

[10] *See Cantor & Hamilton (2007)*
[11] *See, for example, DeRosa-Farag* et al *(1999)*
[12] *For example, the cohort of B-rated issuers formed on January 1, 1996 contained 519 firms. Ten years later, there were just 84 surviving issuers left from this cohort*
[13] *In figure 3, we show as expected that the bootstrap estimates of marginal default rates are less precise as the time horizons are lengthened, particularly for speculative-grade companies. (The marginal default rate in year t is the probability of defaulting in year t conditional on having not defaulted through year t–1, whereas the cumulative default rate in year t is the probability of defaulting during any year before or during year t)*
[14] *See Cantor & Mann (2006)*

4 Five-year cumulative default rates for monthly investment-grade cohorts



5 Five-year cumulative default rates for monthly speculative-grade cohorts



formed at different monthly intervals.[15] As is evident, default rates vary over time; however, to a very large extent, the expected rank ordering of risk is maintained across rating categories for each monthly cohort.[16] Although the variation in realised default rates across cohorts formed at different points in time has been very large, Moody's ratings have provided consistent and powerful rank orderings of default risk within individual cohorts. For example, since 1992, there has not been a single instance in which a higher-rated cohort has experienced a higher five-year cumulative default rate than a lower-rated cohort that was formed on the same month.

Conclusion

Default rates are key parameters used in various credit risk management and rating applications. In addition to estimates of long-run average default rates for a rating class, estimates of the standard error of the estimates are also important. In this brief article, we have used bootstrap methods to derive such standard error for the mean cumulative and marginal defaults associated with different rating categories and different horizons. The bootstrap method is a straightforward (though computationally intensive) method for measuring this type of uncertainty.

Our analysis of Moody's data reveals several interesting findings:

- Bootstrap estimates of the standard errors of mean cumulative default rates are relatively tight for lower rating categories and at longer horizons.
- Even for the low-default portfolio portion of the rating scale (Aaa, Aa and A), the estimated mean default rates are significantly different from one another at the two-year horizon.
- While our measures of the long-run average cumulative default rates are in many cases fairly precisely estimated, we should not expect that any individual cohort will necessarily perform very close to the mean. ■

Richard Cantor is managing director, credit policy at Moody's Investors Service in New York. David Hamilton is senior director, credit strategy research at Moody's Analytics in New York. Jennifer Tennant is an analyst, credit policy at Moody's Investors Service in New York. Email: richard.cantor@moodys.com, david.hamilton@moodys.com, jennifer.tennant@moodys.com

[15] *For ease of presentation, we have omitted the Caa default rates from figure 5, which exhibit high volatility at some time horizons due to small sample sizes*

[16] *It is important to note that the historical variation in cohort CDRs has two causes, one being changes in macroeconomic conditions (the 'underlying' default rate) and the other being random chance due to the finite sample size of any individual cohort. When the samples are large, which is common in the later years of our data sample, most of the default rate fluctuations are clearly attributable to changes in the underlying risk of default rather than chance realisations within the sample. Separating these two effects formally, however, would require adding more structure to the analysis*

References

Basel Committee on Banking Supervision, 2005
Validation of low-default portfolios in the Basel II framework
Basel Committee Newsletter 6, September

Cantor R and E Falkenstein, 2001
Testing for rating consistency in annual default rates
Journal of Fixed Income, September, pages 31–56

Cantor R and D Hamilton, 2007
Adjusting corporate default rates for rating withdrawals
Journal of Credit Risk 3(2), pages 3–25

Cantor R and C Mann, 2006
Analyzing the tradeoff between ratings accuracy and stability
Journal of Fixed Income, spring, pages 60–68

Christensen J, E Hansen and D Lando, 2004
Confidence sets for continuous-time rating transition probabilities
Journal of Banking & Finance 28, pages 2,575–2,602

DeRosa-Farag S, J Blau, P Matousek and I Chandra, 1999
Default rates in the high yield market
Journal of Fixed Income, June

Hamilton D, 2007
Default and recovery rates of corporate bond issuers, 1920-2006
Moody's Global Credit Research, February

Hamilton D and R Cantor, 2004
Rating transition and default rates conditioned on outlooks
Journal of Fixed Income, September, pages 54–70

Hanson S and T Schuermann, 2006
Confidence intervals for probabilities of default
Journal of Banking & Finance 30, August, pages 2,281–2,301

Stein R, 2006
Are the probabilities right? Dependent defaults and the number of observations required to test for default rate accuracy
Journal of Investment Management, first quarter, pages 61–71

Witt G, 2004
Moody's correlated binomial default distribution
Moody's Structured Finance Rating Methodology, August