FDIC Center for Financial Research Working Paper

No. 2009-10

How Well Does the Vasicek-Basel AIRB Model Fit the Data?
Evidence from a Long Time Series of Corporate Credit Rating Data

November 2009

FDIC CFR
Federal Dposit Insurance Corporation • Center for Financial Research

How Well Does the Vasicek-Basel AIRB Model Fit the Data?

Evidence from a Long Time Series of Corporate Credit Rating Data

by

Paul H. Kupiec*
November 2009

ABSTRACT

I develop methods that produce consistent estimates of the Vasicek-Basel IRB (VAIRB) credit risk model parameters. I apply these methods to Moody's data on corporate defaults over the period 1920–2008 and assess the model fit and construct hypothesis tests using bootstrap methods. The results show that the VAIRB does not capture the variability in Moody's default data: there are numerous episodes in which obligors default with much greater frequency than predicted. This pattern is consistent with a missing common factor that affects default correlation only intermittently—a missing factor similar to the frailty covariate in Duffie et al. (2009). Unlike Lopez (2004), I find the VAIRB correlation parameter to be larger for lower-rated credits. I use estimates of the VAIRB error distribution to construct capital allocations for model risk and find that the capital buffers for model risk are substantial, especially for lower-graded credits. VAIRB common factor estimates exhibit positive autocorrelation and thus long time series are usually necessary to produce reliable model estimates. Alternatively, I use common factor and correlation parameter estimates from the 1920-2008 data to control for common factor realizations when estimating unconditional default rates (PDs) from short samples. I estimate PDs and confidence intervals using default data for Moody's alpha-numeric rating grades (1998-2008). After correcting for common factor effects, sample average default rates are shown to overstate the PD for most credit grades in this sample period.

---

* Federal Deposit Insurance Corporation. The views expressed are my own and do not necessarily reflect the views of the FDIC. For comments on a preliminary draft of this paper, I am grateful to Bob Jarrow, Ed Kane, Matt Pritsker, members of the Basel Implementation Validation Subgroup; participants in seminars at the Federal Reserve Bank of Richmond and the Federal Reserve Bank of Cleveland, and discussions at the Federal Reserve Bank of San Francisco January 2009 Conference on Financial Markets, the interagency quantitative risk modeling workshop held at the Federal Reserve Bank of Philadelphia, and the September 2009 Bocconi University CAREFIN conference.

E-mail: pkupiec@fdic.gov.

HOW WELL DOES THE VASICEK-BASEL AIRB MODEL FIT THE DATA?

EVIDENCE FROM A LONG TIME SERIES OF CORPORATE CREDIT RATING DATA

I. INTRODUCTION

The Basel II Advanced Internal Ratings-Based (AIRB) framework was developed to set minimum regulatory capital requirements for the largest and most sophisticated internationally active banks.[1] The Financial Stability Institute (2006) reports that 95 countries plan to implement Basel II by 2015, and more than 60 percent of them plan to include the AIRB option for credit risk capital requirements.

The AIRB regulatory framework uses an asymptotic version of Vasicek's (1987) portfolio credit loss model to approximate the annual default rate distributions on portfolios of credits that are differentiated by a bank-assigned credit rating. To approximate the portfolio credit loss distribution for each credit grade, the AIRB framework uses the Vasicek default rate model along with bank estimates of loss given default (LGD) and exposure at default (EAD).[2] Regulatory capital requirements are set equal to the 99.9 percent upper-tail critical value of the portfolio loss distribution associated with each credit grade.

Much has been written about Basel II, but few if any studies have estimated the model's parameters directly from default rate data alone or have analyzed how well the AIRB model fits a long time series of default data produced by portfolios of credits categorized under a consistent credit rating system.

---

[1] The Basel Committee on Banking Supervision (2006), hereafter BCBS.

[2] The LGD and EAD estimates are not part of the Vasicek model but are calculated independently of the parameters of that model.

This paper develops a new approach for estimating the parameters of the Vasicek-AIRB model (VAIRB). Panel regression methods are used to construct VAIRB parameter estimates from time series data on a cross section of failure rates from a consistent credit rating system. The new approach differs from the ususal "calibration" approach in which the model's correlation parameter is estimated using stock return data (e.g., Zeng and Zhang [2001], Lopez [2004]) and unconditional loss rates are inferred from a separate analysis of default data. My new proposed method estimates all VAIRB parameters simultaneously using only data on observed default rates. The methodology produces consistent estimates of the unconditional probability of default (PD) associated with each credit grade, the VAIRB default correlation parameter, and the common factor realizations that drive default correlation.[3]

I implement the estimation methodology using default rate data from Moody's Investors Service on rated corporate bond issues over the period 1920–2008. I construct consistent estimates and the sampling distributions for PDs associated with Aa, A, Baa, Ba, B, and CaaC rating grades as well as for the VAIRB default correlation parameter. Moody's data provide perhaps the longest times series of default rates for specific credit grades assigned under a systematic credit-rating system. Moody's data were used as a reference when the AIRB framework was developed and are implicitly recognized by the Basel Committee on Bank Supervision as an acceptable benchmark of comparison for calibrating the PDs associated with a bank's internal ratings system.[4]

---

[3] The methodology can accommodate multiple correlation parameters if the data include a sufficiently large number of credit grades.

[4] This rating agency "mapping" approach is described in BCBS (2006), p. 102, paragraph 461-463.

The results show that the VAIRB is unable to accurately reproduce the observed variation in Moody's corporate default data. Observed default rates among credit classes are not as highly correlated as the VAIRB predicts. Moreover, within each grade there are episodes when credits default with far greater frequency than is predicted by the model.

The empirical shortcomings of the VAIRB are not unexpected given the findings of related studies of default correlations.[5] For example, using a reduced form doubly-stochastic default intensity model, Duffie, Eckner, Horel, and Saita (2009) find that a model with a common latent factor and observable macroeconomic variables is incapable of controlling for the correlation in firm-level corporate default intensities. Duffie et al. model the additional default correlation with an unobserved time-varying covariate that they term a "frailty" factor. The frailty factor can generate increases or decreases in default correlation, but it does so only intermittently. By comparison, the VAIRB is a simple and restrictive specification, so its inability to replicate the default patterns observed in the data is unsurprising. Still, given the importance of the VAIRB, including its use to set international bank capital requirements, it is important to quantify the model's accuracy regardless.

Formal hypothesis tests are constructed to investigate whether the default rate patterns in the data are consistent with a single VAIRB correlation parameter. The test results show that different correlation parameters are needed to model high-quality and low-quality credit grades. In particular, Moody's lower-quality credit grades require a much larger correlation parameter to explain observed default rates. This result

[5] These include, *inter alia*, Das, Duffie, Kapadia, Saita (2007), Das, Freed, Geng, and Kapadia (2006), and Duffie, Eckner, Horel, and Saita (2009).

contradicts the findings of Lopez (2004), who used stock return data and KMV-style factor models (Zeng and Zhang [2001]) to indirectly infer VAIRB default correlation parameter values. Lopez concluded that higher-quality credits require a larger correlation parameter, and these findings were used by the BCBS to calibrate the Basel II AIRB model.[6]

At present, there is no widely-accepted robust method for verifying the accuracy of credit grade PD estimates derived from small samples. The VAIRB specifies that default rate realizations are driven by a common factor which should be accounted for when PDs are estimated. The data, moreover, indicate that default rates have strong positive autocorrelation. This autocorrelation makes it impossible to reliably estimate PDs from small samples unless there is a control for the common factor. Consistent estimates of the VAIRB common factor realizations and correlation parameter can be recovered from long time series panel data and can be used to construct consistent estimates of the PD and associated confidence interval of a new credit grade even when that grade has only a limited sample history. I derive the algorithm to correct for common factor realizations when estimating PDs and PD sampling distributions. I demonstrate the procedure using default data on Moody's alpha-numeric rating scale over the period 1998–2008. Over this sample, which includes 7 years of above-average default rate experience, the common factor correction reduces most credit grade PD estimates relative to the sample average PD estimator. In all but a few cases involving low default rate portfolios, the common factor correction produces a tighter sampling distribution for the

[6] The AIRB capital rule includes a regulatory correlation function for corporate credits that mimics the findings of the Lopez study.

PD estimate relative to the sampling distribution associated with the uncorrected sample average default rate estimate of PD.

The proposed method for controlling for the latent factor realizations when estimating PD and PD confidence intervals is novel in the literature.[7] Some studies have attempted to account for common market effects on PD estimators by adjusting confidence interval estimates, but none has directly controlled for the common effects specified by the VAIRB.[8]

Under the VAIRB model assumptions, zero default rate observations (ZDROs) should almost never occur, and yet one-third of the Moody's sample are ZDROs. The unexpectedly high frequency of ZDROs may imply that the Vasicek default rate model is inappropriate for the data, but it may also be a consequence of measurement error. The measurement error arises because the data are not generated by the pure asymptotic portfolios modeled by the VAIRB.

Lando and Skødeberg (2002) develop a continuous time model that uses ratings transition data and Markov-chain methods to impute default rates for rating classes with few or no recorded defaults over short (one-year) horizons. This approach uses data on the length of time (the duration) each credit remains in a rating grade. The VAIRB estimation approach proposed here uses cohort data and so the Lando and Skødeberg method cannot be applied. Pluto and Tasche (2005) suggest an alternative method for

---

[7] The typical approach for estimating one-year unconditional default rates using cohort data does not control for year (common factor) effects. Data are pooled for multiple years, and unconditional default rates are estimated as the sample proportion of credits in a rating grade that default within a one-year horizon.

[8] These studies include Cantor and Falkenstein (2001), Hanson and Schuermann (2006), and Cantor, Hamilton, and Tennant (2007).

deriving confidence bands for PDs under the maintained assumption that credit grades only weakly rank-order obligors. Their method is also inappropriate for purposes of this paper as it produces confidence bands for PDs, not point estimates, and it is not applicable when many of the rating grades have measureable default rates.

I treat ZDROs as observations with measurement error and consider alternative upper bounds on the magnitude of this error. I report VAIRB parameter estimates when ZDROs are truncated to alternative values. The results show that VAIRB parameter estimates are sensitive to the treatment accorded ZDROs. While the approach recognizes the importance of measurement error in generating ZDROs, it does not address the problems created by measurement error in positive default rate observations. There are many unsettled questions that could benefit from additional research related to statistical inference from cohort default data that includes measurement error.

The explanatory power of the VAIRB has important implications for the prudential regulation of bank minimum capital requirements. The VAIRB is the basis for international minimum capital regulations for the largest and most complex banking institutions, yet the model has only a weak ability to reproduce spikes in observed default rate data. To assess the importance of this shortcoming, I estimate the capital that would be needed as a buffer against VAIRB model prediction error for each of the Moody's credit grades. The implied capital requirements for VARB model risk are economically important for all credit grades, and for lower-quality credits, model risk capital far exceeds the credit risk capital assigned by the Basel II AIRB rule.

The next section reviews the VAIRB portfolio default rate model. Section III describes the proposed method for estimating the VAIRB model parameters. Section IV discusses VAIRB parameter estimation when default rates are autocorrelated. The Moody's corporate default rate data are discussed in Section V, and Section VI discusses the interpretation and treatment of ZDROs. Section VII reports model parameter estimates and hypothesis test statistics. Section VIII analyzes the implications of VAIRB prediction errors for internal capital allocation and for the adequacy of Basel II AIRB regulatory minimum capital levels. Section IX provides evidence on the magnitude of bias in small-sample estimates of PDs. Section X develops the correction algorithm that controls for the autocorrelation of common factor realizations in a small sample estimator for the PD of a credit grade and applies the procedure to default rate data for credits rated using Moody's alpha-numeric rating scale. A final section summarizes the results and concludes. A short appendix demonstrates the consistency of the proposed VAIRB estimators.

II. The Vasicek Portfolio Default Rate Model

The Gaussian single factor model of portfolio credit losses developed by Vasicek (1987), Finger (1999), Schönbucher (2001), Gordy (2003), and others provides an approximation for the distribution of the default rate on a well-diversified credit portfolio. The asymptotic version of the Vasicek model that is used in the Basel II AIRB model (VAIRB) focuses on a large diversified portfolio in which idiosyncratic risk is fully diversified and the only source of portfolio default rate uncertainty is the realization of a single common latent Gaussian factor.[9]

[9] The VAIRB assumes that the PD, exposure at default, and loss rates in default (LGD) are known nonstochastic quantities for all obligors.

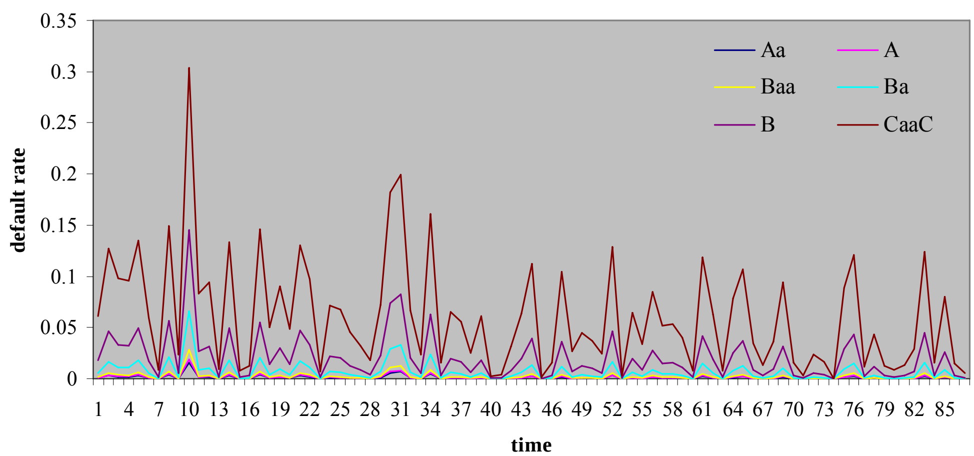
Default on credit i is triggered by the realization of a latent unobserved factor, $\widetilde{V}_i$, which is interpreted as a proxy for the asset value of the firm that issued credit i. $\widetilde{V}_i$ is determined by two random Gaussian factors, one of which, $\widetilde{e}_M$, is common to the latent factors associated with each credit. $\widetilde{V}_i$ is distributed standard normal, and credit i is assumed to default when its latent factor realizes a value less than a credit-specific threshold, $\widetilde{V}_i < D_i$. The unconditional probability that credit i defaults is $PD_i = \Phi(D_i)$, where $\Phi(\cdot)$ represents the cumulative standard normal density function. The common factor, $\widetilde{e}_M$, induces correlation between individual credit latent factor realizations, $\rho = \frac{Cov(\widetilde{V}_i, \widetilde{V}_j)}{\sigma(\widetilde{V}_i)\sigma(\widetilde{V}_j)}$, which induces a correlation among default realizations.

An asymptotic portfolio includes an infinite number of individual credits with identical PD and default correlation parameters. In an asymptotic portfolio, idiosyncratic risks are completely diversified, and therefore the portfolio's realized default rate is driven by the common factor alone. The default rate for an asymptotic portfolio composed of credits with a PD of default PD_i and a default correlation parameter ρ is given by $\widetilde{X}_i$,

$$\widetilde{X}_i = \Phi\left(\frac{\Phi^{-1}(PD_i) - \sqrt{\rho}\ \widetilde{e}_M}{\sqrt{1-\rho}}\right). \qquad (1)$$

Exhibit 1 shows plots of simulated default rates on six hypothetical asymptotic portfolios (bonds rated Aa, A, Baa, Ba, B, and CaaC) where credit defaults are determined by the VAIRB. The correlation parameter is assumed to be 0.20, which is typical of the correlation value used in many applications for corporate credits. The simulated default rates on the portfolios in Exhibit 1 are very highly correlated because the idiosyncratic risk of default is completely diversified and the default rate is driven by only a single common factor. Under the VAIRB, realized portfolio default rates will be nearly perfectly correlated for all non-zero ρ values. The default rates would be exactly perfectly correlated except that equation (1) applies different nonlinear transformations to the common Gaussian term e_M. In Figure 1, the sample pairwise correlations are in excess of 0.98.

Exhibit 1: Simulated Time Series of Vasicek Model Default Rates



The simulations are based on the asymptotic Vasicek model with unconditional default rate values of 15 bps (Aa), 18 bps (A), 29 bps (Baa), 82 bps (Ba), 229 bps (B) , 655 bps (CaaC) and a correlation parameter of 0.20.

III. ESTIMATION OF THE ASYMPTOTIC VASICEK MODEL PARAMETERS

I adopt the common practice of identifying a credit rating (credit grade) with its constant unconditional probability of default. Let X_{jt} represent the realized default rate on an asymptotic portfolio associated with credit grade *j* in year *t*. Equation (1) implies

$$\Phi^{-1}(X_{jt}) = \frac{\Phi^{-1}\left(PD_j\right)}{\sqrt{1-\rho}} - \frac{\sqrt{\rho}}{\sqrt{1-\rho}} e_{Mt} . \tag{2}$$

Observed default rates may deviate from their theoretically predicted value by a mean-zero error term, $\tilde{\varepsilon}_{it}$. Consistent with the VAIRB model assumptions, the error terms for a rating grade are assumed to be independent and identically distributed across time. Rating grade errors are assumed uncorrelated within a cross section, but each rating grade may have its own residual variance,

$$\begin{aligned}
&E\left(\varepsilon_{it}\right) = 0 \quad \forall\ i,t \\
&E(\tilde{\varepsilon}_{it}\, \tilde{\varepsilon}_{jt}) = 0 \quad \forall\ j \neq i \\
&E\left(\tilde{\varepsilon}_{it}\, \tilde{\varepsilon}_{jt+k}\right) = 0,\ for\ k = 1,2,3,\ldots,\ and\ \forall\ j, \\
&E(\tilde{\varepsilon}_{it}\, \tilde{\varepsilon}_{jt}) = 0 \quad \forall\ j \neq i \\
&E\left(\tilde{\varepsilon}_{it}{}^{2}\right) = \sigma_i{}^{2} .
\end{aligned} \tag{3}$$

Incorporating model error, the empirical specification of the VAIRB is

$$\Phi^{-1}(\tilde{X}_{jt}) = \frac{\Phi^{-1}\left(PD_A\right)}{\sqrt{1-\rho}} - \frac{\sqrt{\rho}}{\sqrt{1-\rho}} e_{Mt} + \tilde{\varepsilon}_{it} . \tag{4}$$

Consistent Estimation of the Transformed Model Parameters

Using the following definitions,

$$y_{jt} = \Phi^{-1}(X_{jt}) \tag{5}$$

$$a_j = \frac{\Phi^{-1}\left(PD_j\right)}{\sqrt{1-\rho}} \tag{6}$$

$$b_t = -\frac{\sqrt{\rho}}{\sqrt{1-\rho}} e_{Mt} , \tag{7}$$

equation (4) can be written,

$$\tilde{y}_{jt} = a_j + b_t + \tilde{\varepsilon}_{jt} \ . \tag{8}$$

Notice that y_{jt}, the asymptotic portfolio default rate transformed by the inverse normal distribution function, is not bounded between 0 and 1 but is a continuous variable in the range $\pm\infty$. a_j is a constant determined by the default correlation parameter and the PD for credit grade *j*. b_t is a scalar multiple of the common Gaussian factor realization; it is independent of each asymptotic portfolio's credit rating.

Under VAIRB assumptions, the parameters a_j and b_t can be consistently estimated in a panel regression model with *N* cross sections and *T* time period. Let $W_{it} = \begin{pmatrix} D1_{it} & D2_{it} & D3_{it} \ldots DN_{it} \end{pmatrix}$ be a $(1 \times N)$ selection matrix that indicates membership in a specific credit grade. For example, the transformed default rate associated with credit grade 1 has $y_{it} = y_{1t}$ and $W_{1t} = \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 \end{pmatrix}$. Similarly the transformed default rate associated with credit grade 2 has $y_{it} = y_{2t}$ and $W_{2t} = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \end{pmatrix}$. Define $\tau_{it} = (\tau 1_{it} \quad \tau 2_{i2} \quad \ldots \quad \tau T_{it})$ to be a $(1 \times T)$ selection matrix that identifies the year associated with observation y_{it}. For example, when $y_{it} = y_{i1}$, a default rate observation from year 1 on credit class i, $\tau_{i1} = (1 \quad 0 \quad 0 \quad 0 \quad \cdots \quad 0)$; when the observation is from year 3, $\tau_{i3} = (0 \quad 0 \quad 1 \quad 0 \quad \cdots \quad 0)$. When this notation is used, an empirical model for a generic portfolio default rate observation is

$$y_{it} = \begin{pmatrix} a_1 & a_2 & a_3 \end{pmatrix} W_{it}^T + \begin{pmatrix} b_1 & b_2 & b_3 & \cdots & b_T \end{pmatrix} \tau_{it}^T + \tilde{\varepsilon}_{it} \tag{9}$$

where $\tilde{\varepsilon}_{it}$ is the residual term.

In order to identify all the model's parameters, I use the VAIRB assumption that the common Gaussian factor is a standard normal variable. This assumption imposes a restriction that the average time effect is 0. I obtain first-stage consistent parameter estimates by estimating equation (9) under the restriction $\sum_{t=1}^{T} b_t = 0$. I use the first-stage model residuals to produce consistent estimates of the residual variances for each credit rating class, and I generate second-stage estimates by using restricted generalized least squares (GLS).

Consistent Estimation of the VAIRB Parameters

Restricted GLS provides consistent estimates of the parameters in equation (9), but the VAIRB parameter estimates are functions of these estimates. The VAIRB common factor, $\tilde{e}_{Mt}$, has unit standard deviation by assumption, so the variance of the time effect coefficient estimates provides an estimate of $\left(\frac{\rho}{1-\rho}\right)$. Because $\hat{b}_t$ are consistent, $\frac{1}{T}\sum_{t=1}^{T}\hat{b}_t^2$ is a consistent estimator for $\left(\frac{\rho}{1-\rho}\right)$. From this relationship, it is possible to solve for consistent estimators of the remaining VAIRB parameters,

$$\hat{e}_{Mt} = \frac{\hat{b}_t}{\sqrt{\frac{1}{T}\sum_{t=1}^{T}\hat{b}_t}} \quad (T = 1,2,3,\ldots,T), \tag{10}$$

$$\hat{\rho} = \frac{\frac{1}{T}\sum_{t=1}^{T}\hat{b}_t^2}{1+\frac{1}{T}\sum_{t=1}^{T}\hat{b}_t^2}, \text{ and} \tag{11}$$

$$P\hat{D}_i = \Phi\left(\hat{a}_i\sqrt{1-\hat{\rho}}\right), \quad i = 1,2,3,\ldots,N\ . \tag{12}$$

IV. AUTOCORRELATION ISSUES

The VAIRB does not include the possibility of credit cycles; the common Gaussian factor realizations are assumed independent over time. If credit ratings are updated annually and efficiently so that each credit grade has a fixed one-year PD going forward, there should be no autocorrelation in the deviations from the credit grade's unconditional default rate unless there is autocorrelation in the common factor that drives defaults.[10]

Exhibit 2 reports first-order autocorrelation estimates for the annual default rates reported on selected corporate credits rated by Moody's Investors Service. Exhibit 2 also reports autocorrelations for the default rate series after they are transformed using the inverse cumulative normal distribution. Default rates and transformed default rates exhibit positive autocorrelation for every credit grade examined except Aa.

[10] There is a separate literature that investigates whether ratings changes are autocorrelated. The literature suggests that a ratings downgrade increases the probability of a subsequent ratings downgrade where as rating up grades seem to have no effect on ratings transition probabilities. See for example Altman and Kao (1992) or Güttler and Raupach (2009).

Exhibit 2: Credit Cycles in the Realized Default Rates on Rated Corporate Credits

dependent variable	intercept	p-value	lagged dependent variable	p-value	R^2
Aa default rate	0.055	0.005	-0.079	0.470	0.006
A default rate	0.049	0.077	0.412	<.001	0.172
Baa default rate	0.140	0.010	0.452	<.001	0.208
Ba default rate	0.506	0.007	0.511	<.001	0.261
B default rate	1.697	0.002	0.520	<.001	0.270
CaaC default rate	9.477	<.001	0.303	0.005	0.091
Φ^{-1}(Aa default rate)	-3.860	<.001	-0.078	0.491	0.006
Φ^{-1}(A default rate)	-1.170	<.001	0.666	<.001	0.444
Φ^{-1}(Baa default rate)	-1.715	<.001	0.478	<.001	0.231
Φ^{-1}(Ba default rate)	-1.031	<.001	0.628	<.001	0.398
Φ^{-1}(B default rate)	-1.176	<.001	0.494	<.001	0.245
Φ^{-1}(CaaC default rate)	-1.218	<.001	0.314	0.003	0.100

Estimates are based on Moody's Corporate Default Rate Data, 1920-2008. Default rates are the number of defaults in the year following a Moody's rating designation divided by the number of rated credits in a credit grade. Default rates are measured as percentages.

The data show that default rates are strongly autocorrelated and yet the VAIRB model does not recognize this possibility. One can show (see the appendix) that—provided expression (9) is estimated over a sufficiently long time series—even if default rates are autocorrelated because of autocorrelation in the latent common factor, the VAIRB model parameters can be consistently estimated.

The ability to generate reliable parameter estimates of the coefficients in expression (9) depends on the ability to recover an accurate estimate of the unconditional mean of the transformed default rate series. To get a sense of the length of the time series that may be needed to generate accurate estimates of the unconditional mean from expression (9) when default rates are positively autocorrelated, I conduct a Monte Carlo study of the small-sample distributions of the sample mean estimates from two alternative autoregressive processes that are representative of the Moody’s corporate default rate dynamics: (i) Baa-rated credits, and (ii) Ba credits.[11] I simulate the respective processes 100 times, with 121 observations in each sample. The first 21 observations are omitted to remove the effect of initial conditions. For each of the 100 samples, I calculate estimates of the sample mean based on alternative subsample lengths, and I also calculate the characteristics of the respective sampling distributions.

Exhibit 3 reports the results of the Monte Carlo analysis. The results show that the sample mean estimator converges toward the true unconditional mean of each process, but the rate of convergence is slow. Even with 100 observations in a time series sample, the sample mean estimate still exhibits some bias as well as significant variability. A comparison of the alternative processes shows that convergence is faster for the mean estimate of the Baa process, which has weaker autocorrelation and a smaller standard error associated with its Gaussian innovation. The simulation analysis suggests that, although the VAIRB PD parameter can be consistently estimated, reliable PD estimates can only be constructed from long time series samples. PD estimates from small-samples are likely to be unreliable.

[11] The autoregressive process estimates are reported in Exhibit 2.

Exhibit 3: Sampling Distribution for the Simple Average Estimated from a Sample Generated by Alternative Autoregressive Processes

sample size	10	20	30	40	50	100
	Baa process: R_t=0.141+0.4557 R_{t-1}+e_t, e_t~N(0,0.4191)					
average	0.249	0.257	0.257	0.262	0.263	0.251
std dev	0.207	0.174	0.136	0.121	0.105	0.075
minimum	-0.307	-0.091	-0.185	-0.040	0.050	0.082
maximum	0.712	0.701	0.663	0.641	0.586	0.429
	Ba process: R_t=0.5037+0.5145 R_{t-1}+e_t, e_t~N(0,1.40)					
average	1.092	1.099	1.079	1.103	1.090	1.076
std dev	0.723	0.570	0.465	0.406	0.362	0.305
minimum	-0.379	-0.651	-0.353	-0.092	-0.076	0.367
maximum	2.972	2.437	2.221	2.164	2.046	2.177

Sampling distribution for the simple sample mean of two autoregressive processes based on 100 bootstrap replications of the indicated sample size. The autoregressive processes are the empirical AR (1) models for the Baa and Ba default rate processes with parameter estimates given in Exhibit 2. The true unconditional sample averages for the AR (1) process are: 0.2590 for the Baa default rate process, and 1.0375 for the Ba default rate process. Each bootstrap sample begins after 21 burn-in iterations to attenuate the effects of initial conditions.

V. PORTFOLIO DEFAULT RATE DATA

I estimated the parameters of the VAIRB using annual default rate data on six different credit rating categories for corporate bonds over the period 1920–2008 as reported by Moody's Investors Corporation (2009). For each of the credit rating grades—Aaa, Aa, A, Baa, Ba, B, and CaaC—Moody's publishes annual default rate performance data.[12] Moody's calculates default rates as the ratio of the number of issuers that were in a credit grade at the beginning of a year but defaulted within the year to the number of issuers that were in the credit grade at the beginning of the year.[13]

[12] The Aaa-rating grade is excluded from the analysis because credits in this grade did not default within the first year and so this cohort data provide no information on the 1-year unconditional default rate associated with an Aaa rating.

[13] Moody's excludes from the default rate calculation all credits with withdrawn ratings, and argues that the exclusion has little effect on the reported default rates.

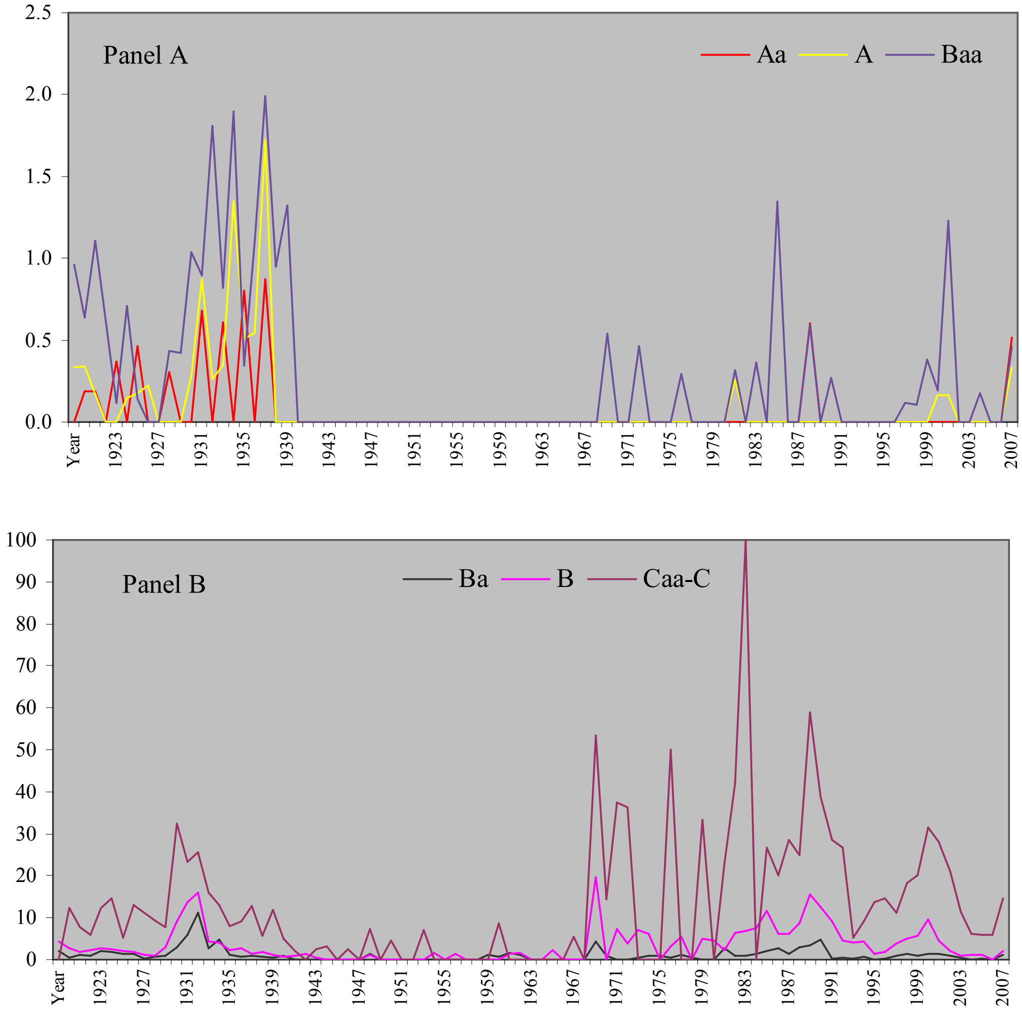
Most rating grades include a large number of bonds in an annual cohort, although some cohorts likely have relatively few bonds.[14] I assume that each annual default rate observation is an approximation for the annual default rate on an asymptotic portfolio of credits, and I interpret a Moody's credit rating as an indicator of the issue's unconditional probability of default.[15] The obligors in a credit grade are assumed to have the same PD over an annual horizon, and this PD is assumed to be fixed over the sample period.

The Moody's annual default rates for credits rated Aa, A, Baa, Ba, B, and CaaC are plotted in Exhibit 4 in two separate panels to accommodate differences in default rate scales.

[14] Moody's does not disclose the number of bonds in each rating grade and cohort for the entire sample period. Rather, it provides partial information on the number of bonds in a rating grade, and these data indicated that in some sample years, the CaaC grade included relatively few bonds.

[15] Moody's argues that a credit rating reflects an assessment of the expected performance of an issue along multiple (unspecified) dimensions and does not represent a ranking based only on the probability of default over a fixed horizon. This claim notwithstanding, it is common to interpret a credit agency rating as an implicit estimate of an issue's probability of default.

Exhibit 4: Moody's Corporate Issuer-rated Default Rates: 1920-2008



The plots in Exhibit 4 show that the default rates on these Moody's rating classes are positively correlated. Exhibit 5 reports the sample correlation estimates along with the sample average annual default rates. The sample correlations reported in Exhibit 5 are not nearly as strong as the sample correlations implied by the VAIRB.

Exhibit 5: Correlation among Moody's Corporate Bond Annual Default Rates for Alternative Ratings Grades, 1920-2008

	Aa	A	Baa	Ba	B	CaaC
Aa	1	0.587	0.357	0.221	0.083	0.022
A		1	0.697	0.363	0.113	0.039
Baa			1	0.556	0.331	0.212
Ba				1	0.711	0.350
B					1	0.642
CaaC						1
average default rate (%)	0.063	0.092	0.271	1.063	3.395	13.103

VI. SPECIAL FEATURES OF THE DATA

There are a number of features of the data that merit discussion. The plots in Exhibit 4 show many observations with zero annual default rates, including 12 years of data for which there are no recorded defaults in any of the credit rating grades.[16] Under VAIRB assumptions, there is virtually no probability that an asymptotic portfolio with a positive PD should experience zero defaults, yet in almost 14 percent of the sample years there are no recorded defaults on any obligor rated by Moody's. Similarly, a 100 percent default rate should be an extremely rare occurrence and, according to VAIRB assumptions, such an extreme default rate must coincide with very high default rates on all portfolios contemporaneously. The Moody's data does not exhibit this pattern.

The prevalence of zeros in the Moody's data (as well as the 100 percent default rate reported for CaaC credits in 1984) can be consistent with the VAIRB if the Moody's rating grade portfolios are not truly asymptotic portfolios—and surely they are not. The

[16] The 12 years when there are no recorded defaults in any of the credit rating grades are 1946, 1948, 1950, 1952, 1953, 1956, 1958, 1959, 1964, 1965, 1967, and 1969. Almost 30 percent of sample are ZDROs.

number of credits in each rating class is limited; accordingly, the observed default rates include measurement error.

Consider a portfolio of 1,000 independent obligors in a single credit grade that did not experience a default within a year. This portfolio is not asymptotic even though it is well diversified by any practical standard. Consider the measured default rate on this portfolio when we add a single credit and the new credit subsequently defaults. This experiment provides the upper bound on this portfolio's default rate of about 10 basis points. Although the observed default rate is zero, the true unobserved default rate could be as large as 10 basis points, given the portfolio characteristics.

Exhibit 6 illustrates the relationship between the number of independent obligors in a credit grade and the magnitude of the upper bound on the potential measurement error associated with a ZDRO. True asymptotic default rates of zero may have zero probability and yet ZDROs may occur simply because of measurement error.

Exhibit 6: Potential Measurement Error and Portfolio Size

number of obligors in a credit grade	upper bound on the magnitude of measurement error
10000	1 bps
2000	5 bps
1000	10 bps
500	20 bps
200	50 bps
100	100 bps

Default rates of zero are also problematic for purposes of estimation. The inverse normal transformation will not accommodate default rates of zero (-∞) or 100 percent (+∞), so these extreme default rate observations must be truncated for estimation. There are many reported default rates of zero in the sample, so the truncation value assigned to

ZDROs will have a measurable effect on the model estimates. I report the estimation results using truncation values for ZDROs.

Measurement error is feature of all observations, not just ZDROs. Although ZDROs and 100 percent default rates are special because their measurement error is necessarily one-sided, a positive default rate observation may also include significant measurement error. When default rates are positive, the measurement error distribution is two-sided but it is also asymmetric.[17] It is an open question how the measurement error in positive default rate observations can best be handled econometrically. In the analysis that follows, I do not account for measurement error in positive default rate observations.

In contrast to the ZDROs, there is only one default rate of 100 percent in the sample.[18] This observation is also likely to be contaminated by measurement error. Again, the measurement error is necessarily one-sided, but unlike ZDROs, this time— since there is only one observation— the truncation value selected has little effect on the results I report. I truncate this observation using the rule 100 percent minus the lower bound used to truncate ZDROs.[19]

Another important choice is the assumption regarding the treatment of the 12 years of data for which there are no observed defaults in any rating category. These observations represent years when there was a very strong economy (i.e., a large positive

[17] Repeat the experiment of adding a single obligor to a large portfolio— but not an asymptotic portfolio. If the obligor subsequently defaults, the default rate changes in magnitude according to the number of obligors in the portfolio. In contrast, if the additional credit does not default, the new portfolio will have a smaller default rate, but the magnitude of the decline will be relatively small.

[18] The default rate reported for the CaaC grade in 1984 is 100 percent.

[19] For example, if ZDROs are truncated to .0001, then the single 100 percent default rate observation is truncated to 1–.0001, or 99.99 percent.

draw from the common Gaussian factor), but the data are not informative as to the magnitude of the positive common factor shock. These data also cannot identify PD differences among the credit grades. These data are included in the estimation sample with an additional restriction that the time dummy variable takes on the same value for each of these years (since they are all equally "good" according to the data). Including these data with uniformly truncated default rates will alter PD estimates as well as the default correlation parameter estimate.[20]

The VAIRB does not include any time dependence in the Gaussian factor structure, so the omission of these 12 years of data does not cause any dynamic inconsistency in the model. However, estimation using a censored sample imparts an upward bias on PD estimates and a downward bias on the estimates of the model correlation parameter. Still, the censored sample estimates are useful for assessing the sensitivity of the VAIRB parameter estimates to ZDROs so I include these results.

VII. MODEL ESTIMATION AND TESTING

Panels A through D of Exhibit 7 report restricted GLS parameter estimates for the VAIRB under alternative assumptions regarding the lower (and upper) bound on portfolio default rates. In Panels A through C the model coefficient estimates that correspond to credit grade covariates are statistically significant and monotonically increasing (from grade Aa to grade CaaC), a pattern that is expected under the VAIRB if credit grade PDs increase monotonically from ratings Aa to CaaC. This monotonic pattern fails to hold in Panel D, when zero default rates are truncated at 50 bps.

[20] These parameters are determined simultaneously, so it is hard to project how truncation will affect the estimates.

Exhibit 7: VAIRB Estimates Based on Moody's Corporate Bond Rating Annual Performance Data 1920-2008

Moody's rating grade	parameter estimate	standard error	t statistic *	implied unconditional PD in bps
Panel A: 0 default rates truncated to 1 bps				
Aa	-3.581	0.061	-51.14	6.7
A	-3.512	0.058	-50.16	8.3
Baa	-3.270	0.053	-46.7	17.0
Ba	-2.760	0.058	-39.41	67.3
B	-2.311	0.064	-33	192.6
CaaC	-1.811	0.086	-25.86	524.1
ρ	0.198			
Panel B: 0 default rates truncated to 10 bps				
Aa	-3.030	0.043	-71.24	18.8
A	-3.011	0.041	-72.75	19.9
Baa	-2.882	0.039	-74.68	29.3
Ba	-2.548	0.040	-64.13	74.2
B	-2.148	0.046	-46.72	199.7
CaaC	-1.628	0.066	-24.51	598.0
ρ	0.085			
Panel C: 0 default rates truncated to 20 bps				
Aa	-2.844	0.037	-75.99	28.4
A	-2.842	0.037	-76.18	28.7
Baa	-2.751	0.036	-77.49	37.4
Ba	-2.476	0.035	-70.22	80.3
B	-2.093	0.041	-51.21	209.1
CaaC	-1.566	0.061	-25.75	640.0
ρ	0.055			
Panel D: 0 default rates truncated to 50 bps				
Aa	-2.579	0.032	-71.24	55.4
A	-2.600	0.034	-72.75	52.3
Baa	-2.564	0.034	-74.68	57.9
Ba	-2.374	0.032	-64.13	96.9
B	-2.015	0.036	-46.72	236.0
CaaC	-1.477	0.055	-24.51	728.8
ρ	0.030			

Parameter estimates are generalized least squares estimates of equation (9) using 89 years of Moody's annual default rate data. All reported t-test statistics are significanly different from zero at the .0001 level of the test.

The estimates in Panels A, B, and C of Exhibit 7 show, predictably, that as the lower bound on ZDROs is increased, the PD estimates increase for all credit grades; at the same time, the estimated value of the VAIRB correlation parameter declines. The treatment accorded ZDROs can have a substantial effect on the correlation parameter's estimated value. The correlation parameter estimate falls from almost 20 percent when the ZDROs are truncated to 1 basis point, to 5.5 percent when the truncation is 20 basis points. These results highlight the importance of ZDROs for statistical inference.

Exhibit 8 reports VAIRB parameter estimates when the data exclude the 12 years for which there are no defaults recorded in any credit rating class. These estimates also imply the anticipated rank ordering between credit quality and PDs provided the truncation value assigned to ZDROs is less than 50 basis points. The VAIRB correlation parameter varies from a high of 15.8 percent when the upper bound on measurement error is assumed to be 1 basis point, to a low of 2.6 percent when the truncation value for ZDROs is set to 50 basis points. Removing the years with no reported defaults has the largest effect on PD estimates for the low-quality credit grades which increase substantially. This is due partly to a smaller estimate of the correlation parameter which reduces the VAIRB model's ability to reproduce the large default rate observations that sometimes occur for lower-rated credits in the Moody's data.

Exhibit 8: VAIRB Estimates Based on Moody's Corporate Bond Rating Annual Performance Data 1920-2008 Excluding Years with No Rated Bond Defaults

Moody's rating grade	parameter estimate	standard error	t statistic *	implied unconditional PD in bps
Panel A: 0 default rates truncated to 1 bps				
Aa	-3.560	0.016	-57.71	5.4
A	-3.480	0.059	-59.04	7.0
Baa	-3.200	0.057	-56.32	16.6
Ba	-2.610	0.064	-40.8	83.1
B	-2.091	0.069	-30.46	274.9
CaaC	-1.514	0.092	-16.49	823.9
ρ	0.158			
Panel B: 0 default rates truncated to 10 bps				
Aa	-3.021	0.049	-70.38	17.6
A	-2.998	0.042	-72.02	18.8
Baa	-2.849	0.040	-71.02	29.5
Ba	-2.463	0.044	-56.03	86.5
B	-2.001	0.049	-40.91	265.6
CaaC	-1.400	0.070	-20.02	881.3
ρ	0.066			
Panel C: 0 default rates truncated to 20 bps				
Aa	-2.839	0.038	-74.82	27.6
A	-2.836	0.038	-75.00	27.9
Baa	-2.731	0.037	-73.94	38.1
Ba	-2.414	0.039	-61.85	91.7
B	-1.971	0.044	-45.15	270.3
CaaC	-1.361	0.064	-21.27	917.5
ρ	0.045			
Panel D: 0 default rates truncated to 50 bps				
Aa	-2.580	0.033	-77.98	54.5
A	-2.604	0.035	-73.69	50.9
Baa	-2.562	0.036	-71.89	57.4
Ba	-2.343	0.035	-66.93	104.0
B	-1.928	0.038	-50.24	285.6
CaaC	-1.306	0.058	-22.52	987.6
ρ	0.026			

Parameter estimates are generalized least squares estimates of equation (9) using 77 years of Moody's annual default rate data. Years in which there are no defaults among the bonds rated by Moody's are excluded from the estimation sample. All reported t-test statistics are significanly different from zero at the .0001 level of the test.

Moody's (2009) reports the number of issuers in each annual rating category cohort from 1970 to 2008. The CaaC ratings category includes relatively few issuers in a number of years of the sample, so these data may include large measurement errors. To assess the importance of this source of measurement error, I estimate the model excluding the CaaC ratings grade (Exhibit 9). When the CaaC data are excluded, there are additional years in which there are no recorded defaults in the Aa, A, Baa, Ba, or B rating grades.[21]

The estimates reported in Exhibit 9 are not materially different from the full sample estimates reported in Exhibit 7, so I conclude that the measurement error bias introduced by including CaaC credits in the estimation is not of first-order importance. Since there are benefits in having an estimate of the CaaC grade PD parameter, I include the CaaC data in the remaining analysis.

[21] The additional years are 1945, 1947, 1951, 1954, and 1968. The coefficients restrictions on zero default rate years are extended to include these years.

Exhibit 9: VAIRB Estimates Based on Moody's Corporate Bond Rating Annual Performance Data 1920-2008 Excluding CaaC Rating Grade

Moody's rating grade	parameter estimate	standard error	t-statistic	implied unconditional PD in bps
Panel A: 0 default rates truncated to 1 bps				
Aa	-3.581	0.048	-73.93	2.7
A	-3.512	0.045	-78.50	3.4
Baa	-3.270	0.045	-73.21	7.8
Ba	-2.760	0.050	-55.73	37.9
B	-2.311	0.059	-39.41	126.9
ρ	0.165			
Panel B: 0 default rates truncated to 10 bps				
Aa	-3.030	0.030	-99.77	14.8
A	-3.011	0.029	-105.37	15.8
Baa	-2.882	0.028	-102.48	23.6
Ba	-2.548	0.032	-80.89	62.4
B	-2.148	0.041	-53	175.7
ρ	0.059			
Panel C: 0 default rates truncated to 20 bps				
Aa	-2.844	0.025	-112.07	22.2
A	-2.842	0.025	-114.16	22.4
Baa	-2.751	0.025	-112.52	29.7
Ba	-2.476	0.027	-91.54	66.4
B	-2.093	0.036	-58.63	181.5
ρ	0.035			

Parameter estimates are generalized least squares estimates of equation (9) using 89 years of Moody's annual default rate data. All reported t-test statistics are significanly different from zero at the .0001 level of the test.

Standard Errors of VAIRB Parameter Estimates

The VAIRB parameter estimates are nonlinear transformations of the restricted GLS parameter estimates of equation (9), so the standard error of these estimates must be obtained from an auxiliary analysis. I construct Efron (1979) bootstrap sampling distributions for the parameter estimates when ZDROs are truncated to two different values: 10 basis points and 20 basis points.[22] Specifically, I draw 5,000 paired samples (with replacement) from the underlying estimation sample of 89 observations, and for each bootstrap sample, I estimate the restricted GLS model parameters and solve for the implied VAIRB parameters. Repeated application builds the sampling distribution for the parameters of the VAIRB. By using paired draws, sampling both the dependent and the independent variables simultaneously, I preserve heteroskedasticity features of the data which are reflected in the sampling distribution of the parameter estimates.

A bootstrap sample may include multiple observations on any year of data, including observations for which there are no observed default rates for any of the rating categories. In the bootstrap exercise, restrictions are imposed to require identical common factor estimates for all years in the sample for which there are no observed defaults in any of the credit grades. As a consequence, the restriction matrix imposed for model estimation is unique in each of the 5,000 bootstrap replications. The summary statistics for the sampling distributions of the VAIRB parameters are reported in Exhibit 10.

[22] For a useful textbook discussion of the bootstrap, see Cameron, A. Colin, and Pravin K. Trivedia (2005).

Exhibit 10: Sampling Distribution for VAIRB Parameter Estimates based on Moody's Corporate Ratings Data, 1920-2008

	Model parameter estimates when 0 default rates are truncated to 10 basis points, 5000						
	Unconditional Probability of Default Parameters						
	Aa	A	Baa	Ba	B	CaaC	ρ
mean	18.78	19.95	29.38	74.59	201.57	608.45	0.085
median	18.68	19.79	29.08	74.04	199.41	594.94	0.085
mode	17.01	18.59	27.87	68.45	169.73	537.36	0.068
std dev	1.67	2.00	3.78	11.06	32.85	140.87	0.011
maximum	26.19	29.27	44.63	129.29	346.53	1350.87	0.128
quantile							
99	23.02	25.17	39.41	102.86	286.82	1000.48	0.111
95	21.75	23.48	36.01	93.90	259.96	858.70	0.103
90	20.97	22.64	34.33	89.15	245.01	793.24	0.099
10	16.74	17.53	24.75	60.94	161.97	440.04	0.070
5	16.23	16.92	23.64	57.43	150.97	398.88	0.067

	Model parameter estimates when 0 default rates are truncated to 20 basis points, 5000						
	Unconditional Probability of Default Parameters in bps						
	Aa	A	Baa	Ba	B	CaaC	ρ
mean	28.79	29.05	38.02	81.49	211.95	651.41	0.057
median	28.71	28.92	37.78	80.90	209.95	635.93	0.057
mode	25.21	24.69	29.48	64.29	141.94	443.02	0.040
std dev	1.66	1.95	3.65	10.08	30.66	138.39	0.009
maximum	35.29	38.27	54.11	135.14	345.97	1358.20	0.094
quantile							
99	33.08	33.99	47.42	106.98	290.34	1030.53	0.079
95	31.65	32.47	44.37	99.20	265.37	896.76	0.072
90	30.93	31.67	42.77	94.87	252.35	834.62	0.069
10	26.72	26.90	33.53	68.92	174.57	484.39	0.046
5	26.21	26.07	32.45	66.06	164.46	445.64	0.044

Bootstrap sampling distribution esimates based on 5000 replications of paired resampling of Moody's Investors Corporate Bond Default Rate Data, 1920-2008. The model estimation restrictions are dynamically modified to impose an identical macro factor value for all resampled observations for which there are no default rates observed in any rating grade.

For the credit categories Aa through Baa, unconditional default rates are fairly accurately estimated. The standard errors of the unconditional default rate sampling distributions are less than 10 percent of the mean value of the PD estimates. For credit ratings in the Ba to CaaC range, the relative precision of the PD estimates declines. For the lowest-quality credits, CaaC, the standard deviation of the sampling distribution of PD estimator is about 23 percent of the mean value when ZDROs are truncated to 10 basis points or 21 percent when ZDROs are truncated to 20 basis points.

These results differ from those reported in at least two studies whose authors calculate estimates of confidence intervals for unconditional default rates by using the continuous Markov-chain duration approach. Both Hanson and Schuermann (2006) and Cantor, Hamilton, and Tennant (2007) report that, when they analyze the sampling distributions for unconditional default rate estimators, lower-quality credits have larger associated coefficients of variation.

Hypothesis Tests for Rank-Order Among Credit Grades

The bootstrap procedure can be used to generate the sampling distributions for hypothesis tests. For example, one measure of the integrity of a rating system is its ability to order credits according to the magnitude of their PDs. The statistical significance of differences in credit grades' PD estimates can be measured with the use of the bootstrap sampling distribution of the difference between two rating grade PD estimates. Exhibit 11 reports descriptive statistics for the sampling distributions of the differences in unconditional default rate estimates associated with adjacent rating grades.

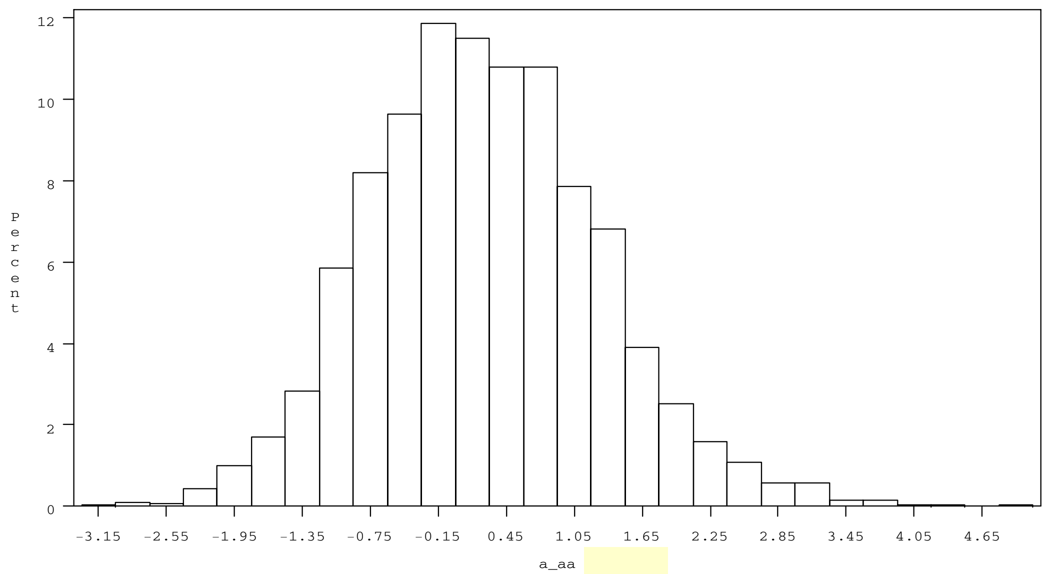
Exhibit 11: Sampling Distribution for Differences in VAIRB PD Estimates based on Moody's Corporate Ratings Data, 1920-2008

Model parameter estimates when 0 default rates are truncated to 10 basis points, 5000 paired sample bootstrap replications					
Difference in Unconditional Probability of Default Parameters					
	A-Aa	Baa-A	Ba-Baa	B- Ba	B-CaaC
mean	1.68	9.43	45.21	126.98	406.88
std dev	1.08	2.43	8.97	26.80	124.09
quantile					
99	4.05	15.88	68.40	197.79	748.07
95	3.05	13.80	60.79	174.52	627.36
5	-0.48	5.81	31.29	86.54	224.59
1	-1.12	4.71	27.25	72.45	171.35
Model parameter estimates when 0 default rates are truncated to 20 basis points, 5000 paired sample bootstrap replications					
Difference in Unconditional Probability of Default Parameters					
	A-Aa	Baa-A	Ba-Baa	B- Ba	B-CaaC
mean	0.27	8.97	43.47	130.46	439.46
std dev	1.01	2.44	8.22	24.99	122.59
quantile					
99	2.98	15.41	64.81	195.02	774.00
95	1.98	13.18	57.81	173.85	658.80
5	-1.28	5.27	30.88	92.39	260.42
1	-1.93	3.99	26.33	78.17	200.42

Bootstrap sampling distribution esimates based on 5000 replication of paired resampling of Moody's Investors Corporate Bond Default Rate Data, 1920-2008. The model estimation restrictions are dynamically modified to impose an identical macro factor value for all resampled observations for which there are no default rates observed in any rating grade.

The results reported in Exhibit 11 suggest that all the rating grades except the adjacent Aa and A categories differentiate credits according to their PD. The results are qualitatively similar irrespective of whether ZDROs are truncated at 10 or 20 basis points. The results suggest that the PD associated with each rating grade increases monotonically from grade A through grade CaaC. However, the sampling distribution of the difference between A- and Aa-rated PDs suggests that these grades do not have statistically different unconditional probabilities of default at conventional levels of significance. Exhibit 12 plots the sampling distribution for the difference between the unconditional default rate estimates for Aa- and A-rated credits based on 5,000 paired bootstrap samples. The sampling distribution clearly highlights the degree of overlap in the Aa- and A-rated sampling distributions.

Exhibit 12: Sampling Distribiution for the Difference Between Aa- and A-Rated PD estimates based on 5000 Bootstrap Replications

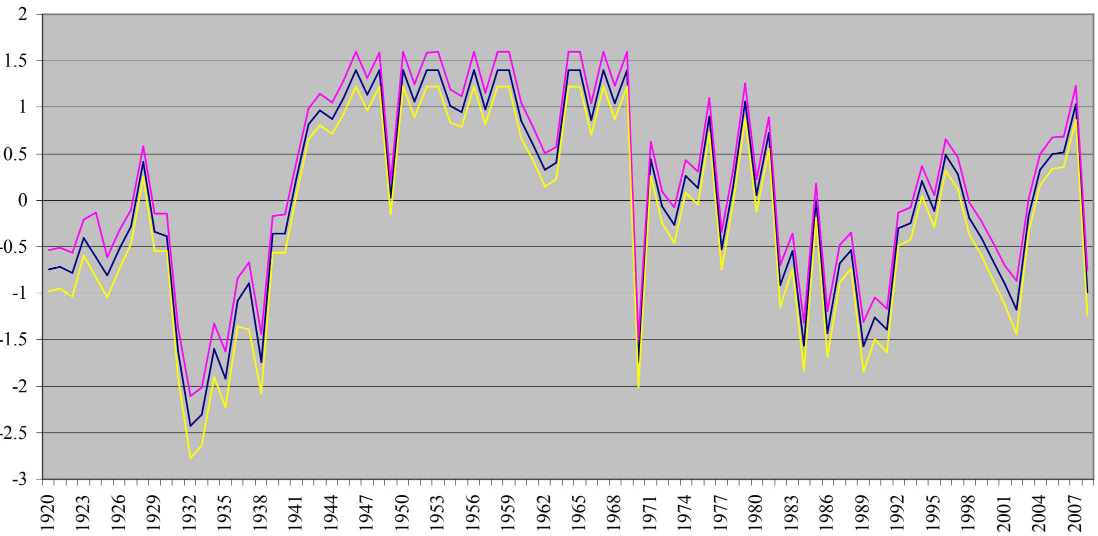


Common Factor Realization Estimates

Exhibit 13 plots the mean and 90 percent probability bounds of the sampling distribution of the VAIRB common factor realizations based on a bootstrap of 5,000 paired replications.[23] Recall that under the VAIRB, positive values of the common factor are associated with low portfolio default rates, whereas negative common factor draws are associated with large default rates.

[23] The 90 percent probability bound comprises the 5 and 95 percentile levels of the common factors' estimated sampling distribution.

Exhibit 13: Sampling Distribution for the VAIRB Common Factor Estimates, 1920-2008



Estimates based on Moody's annual default rate data on rated corporate credits, 1920-2008. ZDROs are truncated to 10 basis points. The sampling distribution is calculated from 5000 paired bootstrap replications. The pink line is the 95th percentile of the sampling distribution. The blue line is the mean of the sampling distribution. The yellow line is the 5th percentile of the sampling distribution.

The common factor estimates plotted in Exhibit 13 show a clear pattern of credit cycles in the default rate data. Realized default rates were above average for most of the 1920s and 1930s and were below average for all but one year over the period 1941 to 1969. This long credit cycle was followed by two shorter credit cycles in addition to the downturn that began in 2008.

The sampling distribution estimates show temporal dependence among the common factor realizations. The Wald and Wolfowitz (1940) runs test statistic for independence is 9.84 when calculated with the use of the mean values of the common factor sampling distributions. The common factor estimates violate a formal nonparametric runs test for independence at any commonly used level of significance.

Thus, the temporal independence assumption of the VAIRB is rejected at normal levels of confidence for these historical Moody's default data.

Test for Multiple Correlation Parameters

The bootstrap can also be used to test other aspects of the VAIRB specification. For example, the restriction of a common correlation parameter across credit rating classes may be questionable. One can construct a formal test by analyzing the sampling distributions for the VAIRB correlation parameter under alternative restrictions.

The baseline specification assumes that the correlation parameter is identical for the entire set of ratings (Aa, A, Baa, Ba, B, and CaaC). Under an alternative restriction, the correlation parameter is estimated and restricted to be identical within the subinvestment-grade group alone (Ba, B, CaaC).[24] Consider the difference between these two restricted correlation estimates. If the correlation is greater for higher-quality credits as assumed in Lopez (2004) and the Basel AIRB framework, then the correlation estimate based on the fully inclusive restriction should exceed the correlation estimate produced when only the more limited (Ba, B, CaaC) restriction is imposed.

The sampling distributions for the alternative correlation parameter estimates and their difference are reported in Exhibit 14. I estimated the sampling distributions by using

---

[24] I construct the test in this manner to avoid losing degrees of freedom. I could have constructed it as a test between separate parameters for the high grade group (Aa, A, Baa) and the low grade group (Ba, B, CaaC), and in that formulation, there would be many new data points for the high-quality group where no defaults were recorded in any of the included credit grades. Constructing the test as I have means that there are no additional years beyond the aforementioned 12 years for which all the rating categories report zero default rates.

5,000 paired-sample bootstrap replications, assuming zero default rate observations are truncated to 20 basis points.

Exhibit 14: Sampling Distributions for Alternative Correlation Parameter Estimates and their Differences

	[1] ρ estimate from Aa, A, Baa, Ba, B, CaaC credits (in pct)	[2] ρ estimate from Ba, B and CaaC credits (in pct)	estimate of difference in correlation [2]-[1]
mean	5.68	18.17	-12.49
median	5.66	18.15	-12.5
mode	6.1	19.57	-13.47
maximum	9.27	24.78	-7.99
99 percentile	7.81	22.42	-9.77
95 percentile	7.13	21.29	-10.59
90 percentile	6.83	20.52	-11.03
10 percentile	4.58	15.82	-13.94
5 percentile	4.29	15.13	-14.38
1 percentile	3.79	13.92	-15.13
minimum	3.05	11.04	-17.02

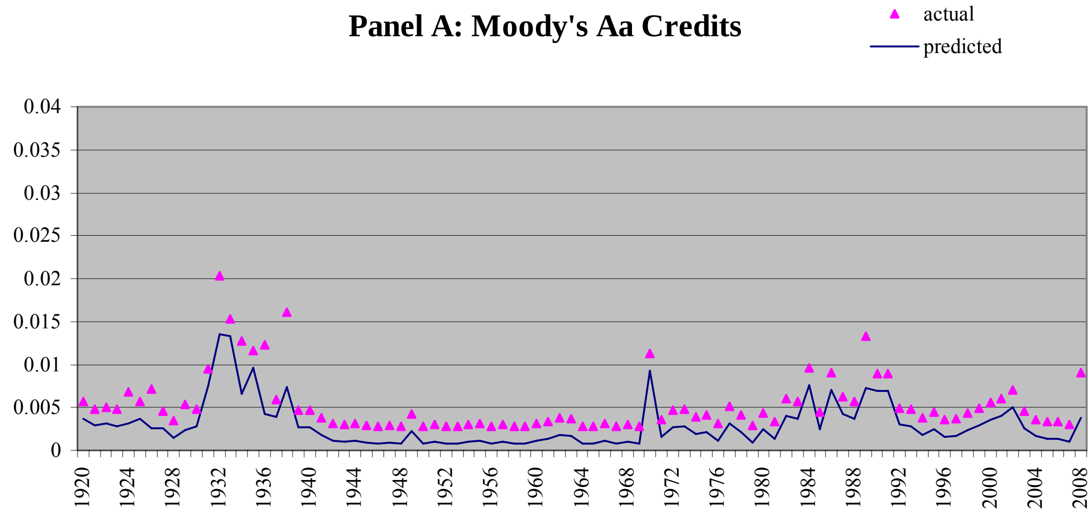
The sampling distributions of the VAIRB model correlation parameter estimates and their differences are estimated using 5000 paired-sample bootstrap replications. In the estimation, zero default rates are truncated to 20 basis points.

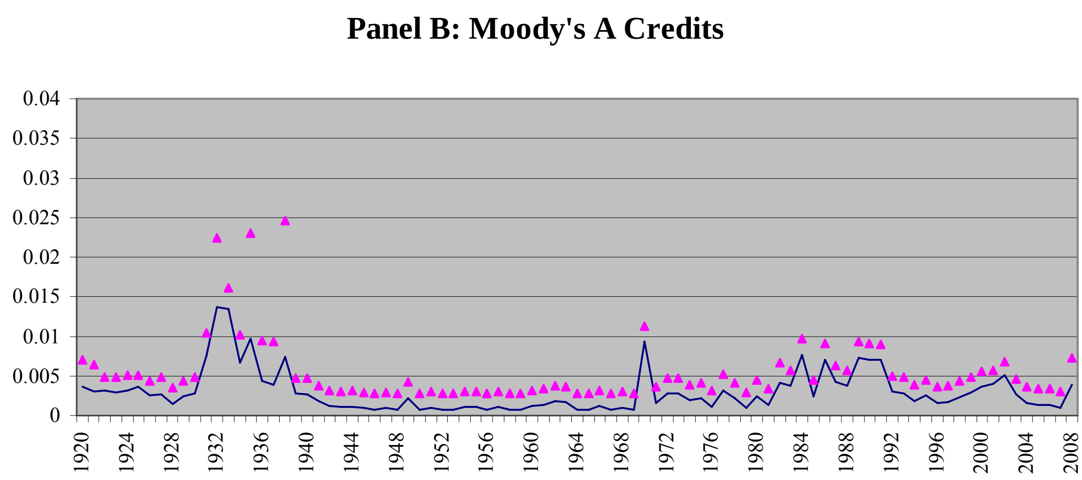
The sampling distribution for the difference in the alternative correlation parameter estimates shows that the Moody's data are consistent with at least two different correlation parameters, one for highly rated credits and another larger correlation parameter for lower-quality credits. This correlation pattern is inconsistent with Lopez (2004) and the Basel AIRB framework. Both suggest that the correlation parameter decreases as a credit's unconditional probability of default increases.

VAIRB Predictive Accuracy

One can assess the fit of the VAIRB by comparing the actual and predicted default rates by credit grade over the estimation sample. Exhibit 15 plots the actual and predicted portfolio default rates for Moody's investment-grade credits. The investment-grade data include a large number of ZDROs. When the investment-grade credit portfolios experience high default rates (in the 1920s and 1930s), the VAIRB produces elevated default rate predictions, but many of these predictions fall far short of the actual recorded default rates. In the period beginning in the late 1960s, the model predicts an elevated level of investment-grade defaults, but the model's default rate predictions are small relative to the actual default rates recorded.

Exhibit 15: Predicted and Actual Default Rates for Moody's Aa-, A-, and Baa-Rated Corporate Credits, 1920-2008, 20 Basis Points Truncation





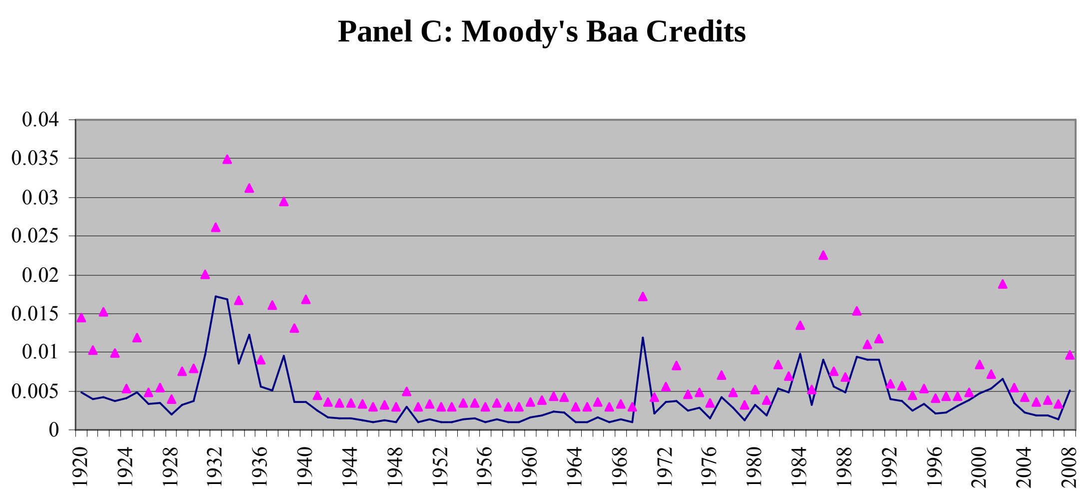
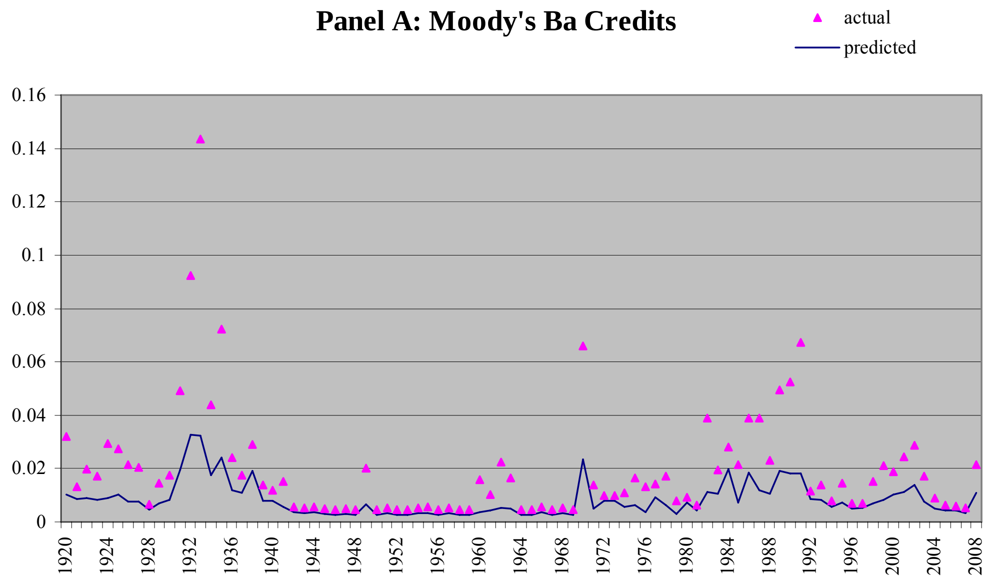


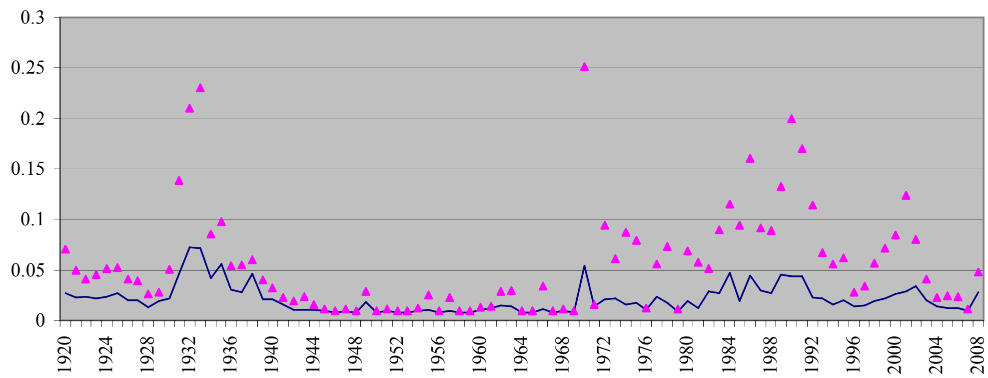
Exhibit 16 provides the actual and predicted portfolio default rates for Moody's subinvestment-grade credits. As the plots show, the fit of the VAIRB is particularly poor subinvestment-grade credits. For the lowest-rated credits, B and CaaC, errors are large and concentrated in two periods: the 1930s and a period that began in the late 1970s. The large error rates for subinvestment-grade credits are likely due at least in part to an inappropriate restriction on this group's correlation parameter. A larger correlation parameter for lower-quality credit grades would provide the VAIRB with additional flexibility to model these extreme default rate observations.

Exhibit 16: Predicted and Actual Default Rates for Moody's Ba-, B-, and CaaC-Rated Corporate Credits, 1920-2008, 20 Basis Points Truncation

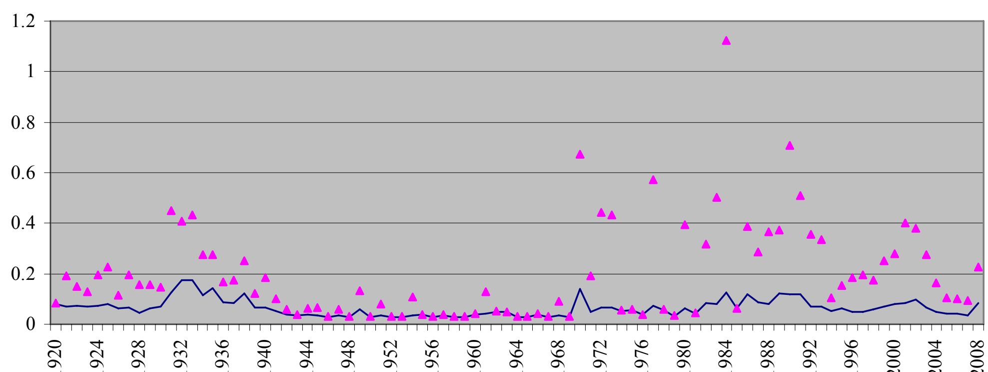
Panel A: Moody's Ba Credits



Panel B: Moody's B Credits



Panel C: Moody's CaaC Credits



The model prediction errors depend partly on the truncation value selected for ZDROs. Within a reasonable range of truncation values, as the truncation value for ZDROs is increased, the unconditional default rate estimates increase and the model correlation parameter estimate decreases. Although I have not conducted an exhaustive analysis of alternative truncation values, a comparison of the RMSE statistics reported in Exhibit 17 suggests that a uniform 20 basis point truncation value is probably a reasonable compromise relative to an objective of minimizing root mean-squared prediction errors across the credit grades.

Exhibit 17: Model Prediction Error Rates for Alternative Zero Default Rate Tuncation Choices

	RMSE for truncation value	
Rating	20 bps	50 bps
Aa	24.4	34.6
A	25.1	35.4
Baa	30.4	36.2
Ba	112.6	112.5
B	328.3	317.9
CaaC	1593.1	1563.7

RMSE is the root mean-square VAIRB prediction error measured in basis points using Moody's Corporate default rate data, 1920-2008.

VIII. VAIRB MODEL ERROR AND ECONOMIC CAPITAL

The magnitude of the VAIRB prediction errors should be a concern for risk managers as well as for bank regulatory authorities. The VAIRB is widely used as a benchmark for setting internal capital allocations for bank investment activities that generate credit risk. Moreover, many countries have adopted or plan to adopt the VAIRB to set minimum regulatory capital requirements for their largest and most complex internationally active banks. The Basel AIRB approach sets minimum regulatory capital requirements equal to the losses associated with the 99.9th percentile of the VAIRB default rate distribution,[25] with an offset for bank reserves held for nonspecific credit losses.[26]

The analysis thus far suggests that the VAIRB is unable to reliably model the default rate patterns that are observed in the Moody's corporate rating data. The results show that, in many circumstances, the VAIRB underpredicts observed default rates, especially when observed default rates are large relative to the sample average experience for a credit grade. These results suggest that there is significant risk that the Basel II AIRB framework may underallocate capital, given its stated objective of covering 99.9 percent of a credit grade's loss distribution.

One solution to this underprediction problem is to include an additional capital component for model risk. In the remainder of this section, I provide estimates of the magnitude of the capital requirements that are necessary to buffer against the model risk generated by VAIRB model prediction error. Economic capital allocations and minimum regulatory capital requirements must include capital for model risk in addition to VAIRB estimates of the capital needed to buffer against credit risk for capital allocations to satisfy regulatory objectives or internal firm objectives.

---

[25] To be precise, capital for unexpected loss is the 99.9 percent default rate from the VAIRB distribution multiplied by EAD and LGD and, in some cases, by a maturity factor. See BCBS (2006), pp. 63 ff. for the exact procedures for calculating minimum regulatory capital requirements for corporate wholesale exposures.

[26] Bank reserves for nonspecific credit losses are a buffer to absorb credit losses. The magnitude of these reserves should approximate so-called expected credit losses and these reserves are subtracted from the unexpected loss estimate. See Kupiec (2003) for additional discussion.

Exhibit 18 reports statistics on the sampling distributions of the VAIRB prediction errors by credit rating. I estimated the sampling distributions from over 31,000 model prediction errors generated by bootstrap resampling. I re-estimated the model for 350 paired-sample bootstrap replications and calculated prediction errors. Zero default rates are truncated to 20 basis points. Exhibit 18 reports that, on average, the VAIRB overpredicts default rates for Aa- and A-rated credits and underpredicts the default rates for the remaining credit classes. The error distributions are strongly skewed, so the mean value is not very informative. The remaining statistics reported in Exhibit 18 include selected percentiles on the upper tails of the VAIRB error distribution (focusing on underprediction).

Exhibit 18: Sampling Distribution Statistics for VAIRB Model Errors by Rating Class

	Moody's Corporate Rating					
	Aa	A	Baa	Ba	B	CaaC
mean	-5.5	-4.5	0.7	30.8	133.1	667.2
median	1.8	0.9	1.9	-3.6	-0.9	165.7
mode	10.8	10.9	8.9	-5.5	-43.3	-138.9
maximum	45.6	111.1	118.8	837.8	1541.4	9115.4
99 percentile	34.3	89.2	99.4	757.3	1373.0	8413.2
95 percentile	13.3	13.2	60.9	195.0	743.2	3430.3
90 percentile	12.6	12.5	28.9	118.9	556.1	2558.4

Model errors are actual minus predicted default rate values. Positive model errors indicate under-prediction. Reported values are measured in basis points. The sampling distributions are estimated from 31,150 residuals generated from 350 paired-sample bootstrap iterations. ZDROs are truncated to 20 basis points.

Exhibit 19 reports the unconditional probability of default estimates and the capital needed to buffer against 99.9 percent of the potential credit loss estimates as calculated by the VAIRB model using Moody's data from 1920-2008 and a truncation value of 20 basis points for ZDROs. In addition to credit risk capital, I calculate the capital needed to buffer an estimate of the 95th percentile of the VAIRB error distribution (from Exhibit 18).[27] The total capital needed to cover 99.9 percent of the estimated credit risk exposure and 95 percent of the estimated model risk exposure is a measure of total capital needed to cover the default risk on these asymptotic credit portfolios.

The estimates in Exhibit 19 show that capital for model error risk contributes less than 10 percent of the total capital needed for the Aa- and A-rated credits. The contribution of model risk to total capital requirements rises as credit quality declines. Model risk comprises about 21 percent of total capital for Baa-rated credits; about 32 percent for Ba-rated credits; and over 50 percent for lower-quality credits. VAIRB model prediction errors are large in many cases, and the results show that there are significant consequences for ignoring model risk when setting minimum capital requirements.

Exhibit 19: Capital Alloacation for Credit Risk and VAIRB Model Error

	Moody's Corporate Rating					
	Aa	**A**	**Baa**	**Ba**	**B**	**CaaC**
Unconditional PD estimate in bps	28.4	28.7	37.4	80.3	209.1	640.0
Capital for 99.9 percent loss coverage in bps*	5000.0	5000.0	5000.0	5000.0	5000.0	5000.0
Capital for 95 percent model error coverage in bps	13.3	13.2	60.9	195.0	7432.0	3430.0
Total required capital in bps	5013.3	5013.2	5060.9	5195.0	12432.0	8430.0
Capital for model error as a percentage of total capital	0.3	0.3	1.2	3.8	59.8	40.7

* The capital estimate is for coverage of 99.9 percent unexpected loss calculated using a correlation parameter estimate of 0.055. The unconditional probability of default and the correlation parameter estimates are based on a truncation rule that sets zero default rates to 20 basis points.

IX. VAIRB PARAMETER ESTIMATION FROM SMALL SAMPLES

Few institutions have data on the default rate performance of their internal rating systems for 89 years. Because of data limitations on banks' own internal ratings system performance, the Basel AIRB approach requires as little as five years of data as the

[27] It would be natural to choose to buffer against 99.9 percent of the model error risk, given regulatory objectives. This calculation, however, is omitted because, in most cases, model risk capital is more than twice as large as the capital needed to buffer 95 percent of the VAIRB model error risk which is already shockingly large. Estimates of the capital needed to buffer against 99.9 percent of the model error risk are reported in Exhibit 18.

minimum acceptable sample length that a bank may use for estimation for some AIRB parameters. An important regulatory concern has been the quality of the parameter estimates that may be generated from such small samples.

Small-sample estimates of unconditional default rate inputs into regulatory capital calculations may be downward biased in samples that do not include sufficient data on a "bust" phase of the credit cycle. To correct for this potential shortcoming, the AIRB requires that the underlying data include either a downturn phase of the credit cycle or, alternatively, some technique to adjust unconditional default rates so that they reflect recession conditions. These so-called "stress PD" requirements are necessary because the AIRB rule does not otherwise include a requirement for banks to control for the effects of the common factor realizations when they are estimating the input values for their minimum regulatory capital calculations.

To assess the reliability of small-sample estimates, I construct sampling distributions for the parameters of the VAIRB when the parameters are estimated from sample sizes of 5 and 10 years of data using the Efron (1979) jackknife re-sampling procedure, with 5,000 paired observations to preserve heteroskedasticity.

Exhibit 20 reports selected statistics on the sampling distributions for VAIRB parameter estimates from sample sizes of 5, 10, and 89 years of data when zero default rates are truncated to 20 basis points. The estimator based on 89 years of data is nearly symmetric, whereas the small-sample distributions are skewed with a long right tail. Because of the skew, small-sample mean estimates are not very informative about the estimates that are likely to be produced in practice. The right tail of the distribution is particularly pronounced when estimates are based on 10 years of data.

Exhibit 20: Small Sample Distributions for VAIRB Parameter Estimates based on Moody's Corporate Ratings Data, 1920-2008

Model parameter estimates based on 10 observations when 0 default rates are truncated to 20 basis points, 5000 paired sample bootstrap replications

	Unconditional Probability of Default Parameters in bps						
	Aa	A	Baa	Ba	B	CaaC	ρ
mean	39.78	40.05	52	107.44	263.7	802.33	0.0773
median	28.97	29.78	38.74	84.51	221.81	688.58	0.0681
mode	35.83	38.22	43.88	100.5	231.81	2136.01	0.0358
maximum	900.67	941.12	1018.02	1412.06	1951.78	5080.93	0.5113
quantiles							
99	212.45	205.16	255.65	485.35	895.2	2617.96	0.2581
95	89.78	90.86	117.95	233.92	545.96	1791.63	0.1651
90	65.03	64.58	85.21	180.09	445.08	1479.88	0.135
10	22.18	22.12	25.68	48.2	116.49	425.6	0.0274
5	20.99	20.91	23.62	41.85	97	263.82	0.02

Model parameter estimates based on 5 observations when 0 default rates are truncated to 20 basis points, 5000 paired sample bootstrap replications

	Unconditional Probability of Default Parameters in bps						
	Aa	A	Baa	Ba	B	CaaC	ρ
mean	29.3	29.76	40.34	90.9	245.46	839.62	0.051
median	26.84	26.37	36.07	77.79	214.02	626.3	0.043
mode	21.01	21.38	24.87	36.06	107.43	51.28	0.011
maximum	383.6	357.52	349.5	704.55	1299.03	6504.09	0.3562
quantiles							
99	61.47	66.82	100.91	269.18	729.21	3813.08	0.1609
95	43.74	47.52	73.12	189.17	528.3	2287.59	0.1191
90	38.95	42.59	62.85	157.74	445.3	1766.05	0.1013
10	21.18	21.11	21.64	39.49	87.74	177.13	0.0114
5	20.56	20.43	20.36	32.23	66.39	118.28	0.0058

Model parameter estimates based on 89 observations when 0 default rates are truncated to 20 basis points, 5000 paired sample replications

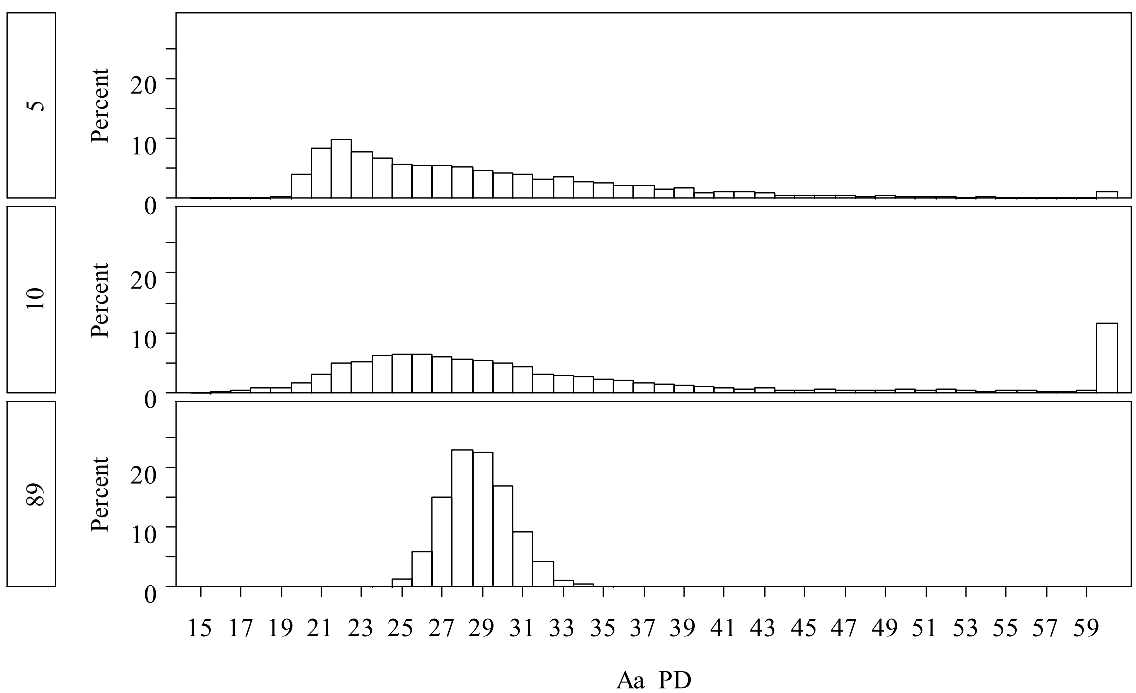
	Unconditional Probability of Default Parameters in bps						
	Aa	A	Baa	Ba	B	CaaC	ρ
mean	28.79	29.05	38.02	81.49	211.95	651.41	0.057
median	28.71	28.92	37.78	80.90	209.95	635.93	0.057
mode	25.21	24.69	29.48	64.29	141.94	443.02	0.040
std dev	1.66	1.95	3.65	10.08	30.66	138.39	0.009
maximum	35.29	38.27	54.11	135.14	345.97	1358.20	0.094
quantiles							
99	33.08	33.99	47.42	106.98	290.34	1030.53	0.079
95	31.65	32.47	44.37	99.20	265.37	896.76	0.072
90	30.93	31.67	42.77	94.87	252.35	834.62	0.069
10	26.72	26.90	33.53	68.92	174.57	484.39	0.046
5	26.21	26.07	32.45	66.06	164.46	445.64	0.044

Jackknife sampling distribution esimates based on 5000 paired resampling of Moody's Investors Corporate Bond Default Rate Data, 1920-2008. The model estimation restrictions are dynamically modified to impose an identical macro factor value for all resampled observations for which there are no default rates observed in any rating grade.

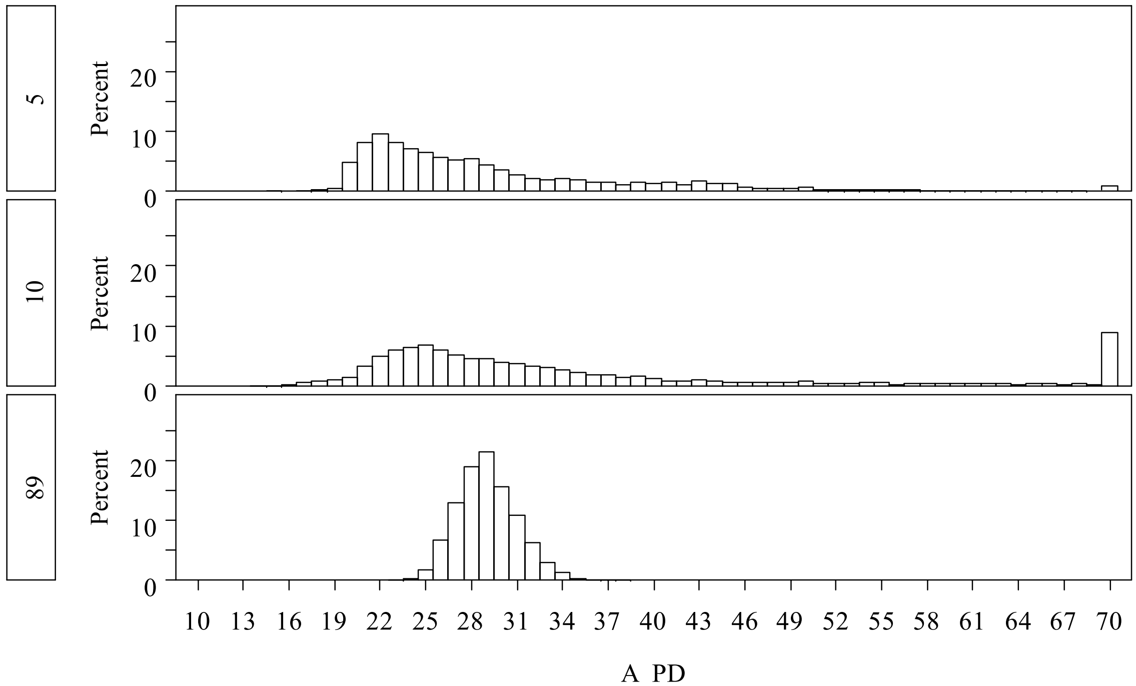
The sampling distributions for VAIRB unconditional default rate parameter estimates are plotted in Exhibit 21. In constructing the histograms, I collected the observations above the largest indicated value on a histogram in the largest bucket included on each histogram. Among these results, relative to the 89-observation sample estimates, there is a pattern in which small-sample parameter distribution estimates have more probability associated with smaller unconditional default rates. Although the right tail of the small-sample estimators puts more mass on very high default rates relative to the estimator based on the 89-observation sample, the total mass on these high default rates is small. Much of the mass of the small-sample estimators is concentrated on smaller default rate values.

Exhibit 21: Small-Sample Distributions of VAIRB PD Estimates

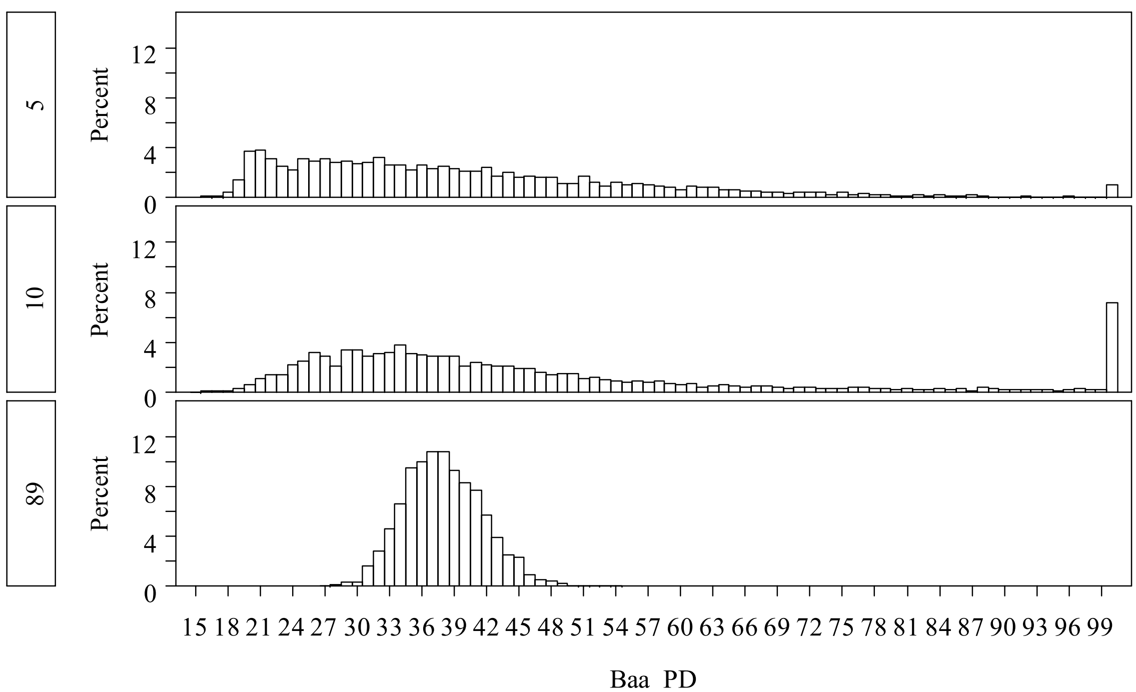
Panel A: Distribution of PD Estimates for Aa-Rated Credits in bps



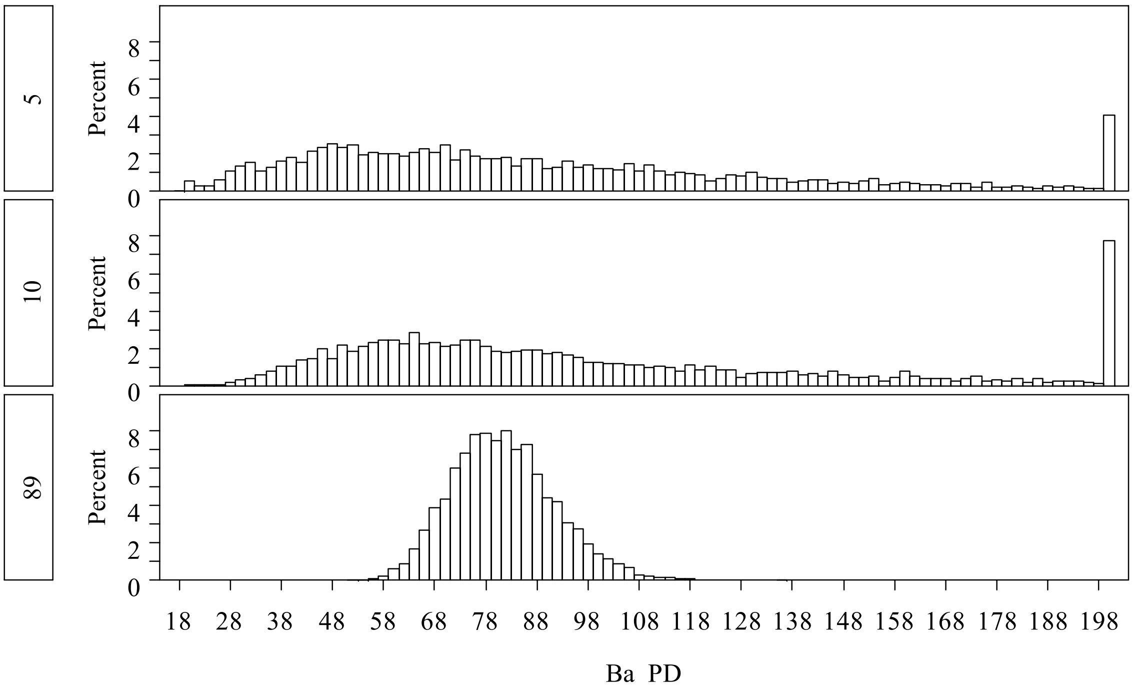
Panel B: Distribution of PD Estimates for A-Rated Credits in bps



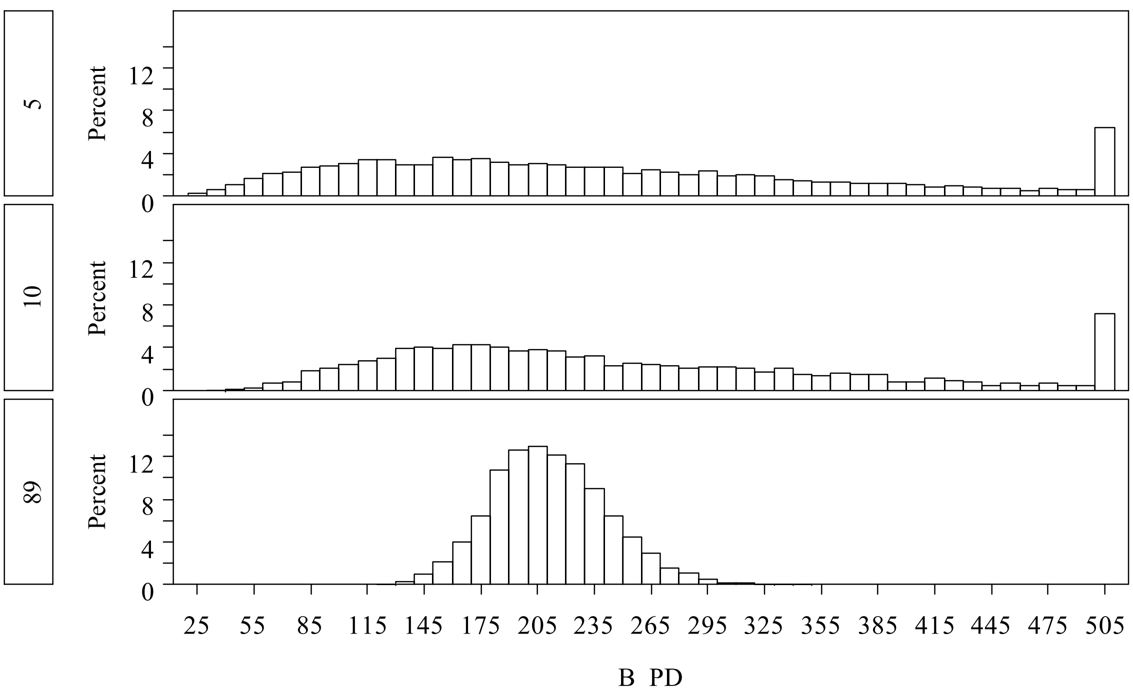
Panel C: Distribution of PD Estimates for Baa-Rated Credits in bps



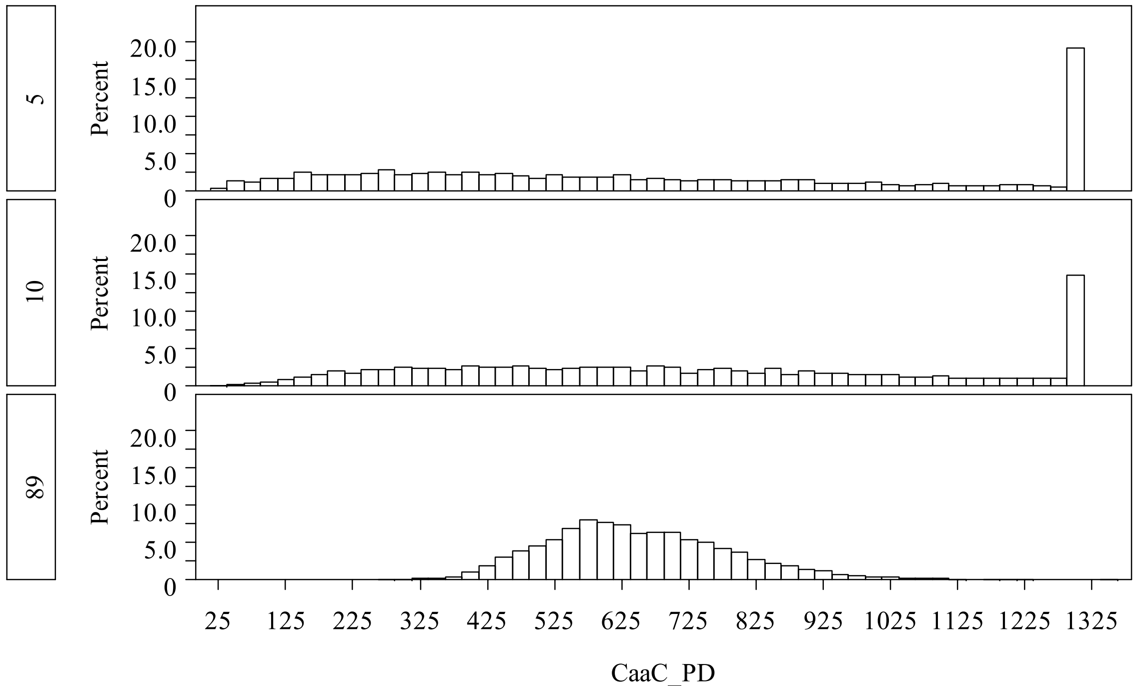
Panel D: Distribution of PD Estimates for Ba-Rated Credits in bps



Panel E: Distribution of PD Estimates for B-Rated Credits in bps



Panel F: Distribution of PD Estimates for CaaC-Rated Credits in bps



There are at least two reasons that the small-sample estimator districutions have more mass concentrated at lower default rates. One reason is the prevalence of ZDROs in the Moody's data. The jackknife random sampling technique will potentially draw a large share of ZDROs in a small sample, given their frequency of occurrence. A second reason for the shape of the distributions is the strong positive autocorrelation in the common factor. Since the common factor is strongly autocorrelated, it will require a very long time series before the common latent factor is likely to have a sample average close to zero. Because VAIRB identification is achieved by imposing the zero mean condition, common factor estimates are likely to be biased in small samples. This bias will also induce a bias in the small-sample unconditional default rate estimates.

The biases that are demonstrated in these small-sample results likely understate the effect of positive autocorrelation in small samples; that is, most of the common factor draws in a small sample are likely to be either positive or negative. In such a case, the unconditional default rate is likely to be poorly estimated. In addition, the model restriction that the common factor averages to zero over the sample will induce a larger bias in the common factor estimates compared with those produced under the jackknife random sampling techniques that underpin Exhibit 21. The small-sample parameter distributions plotted in Exhibit 21 are for samples of observations chosen at random from the entire time series, so the underlying autocorrelation structure in the raw data is not preserved in these jackknife sampling distributions.

X. ALTERNATIVE SMALL-SAMPLE PD ESTIMATE

One can use consistent estimates of the common factor realizations to estimate the unconditional default rates associated with additional credit grades that may not have a long time series of default rate data. The estimator is consistent under the assumption that the default rate and correlation factor associated with the new credit rating class are identical to those that characterize the credit grading system used to produce the common factor estimates.

Recall that the VAIRB model implies

$$\Phi^{-1}(\tilde{P}_{jt}) = \frac{\Phi^{-1}\left(PD_j\right)}{\sqrt{1-\rho}} - \frac{\sqrt{\rho}}{\sqrt{1-\rho}} e_{Mt} + \tilde{\varepsilon}_{it} . \tag{13}$$

The year fixed-effect estimates from (13) are estimates of called common factor realizations, $\hat{b}_t = -\frac{\sqrt{\rho}}{\sqrt{1-\rho}} e_{Mt}$.

Let PD_A represent the unconditional probability of default on an auxiliary rating category for which data are available, but with only a modest sample size of *S*. From expression (13), it is evident that

$$\Phi^{-1}(\widetilde{P}_{At}) - \hat{b}_t = \frac{\Phi^{-1}(PD_A)}{\sqrt{1-\rho}} + \widetilde{\varepsilon}_{it}\,. \tag{14}$$

Consequently, as the size of the small sample for the alternative rating grade increases,

$$S^{-1}\sum_{t=q}^{S}\left(\Phi^{-1}(\widetilde{P}_{At}) - \hat{b}_t\right) \xrightarrow[a.s..]{} \frac{\Phi^{-1}(PD_A)}{\sqrt{1-\rho}}\ . \tag{15}$$

Thus a consistent estimate of the unconditional default rate of the new rating class is

$$\Phi\left(\sqrt{1-\hat{\rho}}\quad S^{-1}\sum_{t=q}^{S}\left(\Phi^{-1}(\widetilde{P}_{jt}) - \hat{b}_t\right)\right) \xrightarrow[a.s.]{} PD_A\,. \tag{16}$$

To estimate the unconditional default rates for Moody's alpha-numeric rating scale, I apply expression (16). Although Moody's publishes default rate statistics on its alpha-numeric rating grades from the early 1980s, I use data from 1998 to 2008 to demonstrate the adjustment. I exclude rating grades that exhibit no defaults over this sample period, and I truncate ZDROs to 20 basis points. I take the correlation and common factor adjustments used in (16) from the VAIRB estimates derived from the Moody's 1920–2008 data on letter rating grades when zero default rates are truncated to 20 basis points (Exhibit 7, Panel C). Estimates of the realization of the common factor estimates for the period 1998-2008 are reported in Exhibit 22. Over this period, there are 4 years in which the common factor realization decreases default rates (2004–2007) and 7 years in which the common factor realization increases default rates (1998-2003, and 2008).

Exhibit 22: Estimates of the Implied Common Factor Realizations from Moody's Corporate Default Data 1988-2008

year	common factor estimate
1998	-0.065
1999	-0.343
2000	-0.634
2001	-0.806
2002	-1.126
2003	-0.206
2004	0.412
2005	0.632
2006	0.646
2007	1.004
2008	-0.745

Estimates are derived from Moody's default data for the period 1920-2008 (See Exhibit 7, Panel C). ZDROs are truncated to 20 bps.

Exhibit 23 reports simple and corrected PD estimates for the Moody’s alpha-numeric sample data. The simple PD estimator is the sample mean. The corrected PD estimator is expression (16). Along with these sample estimates, Exhibit 23 also reports selected characteristics of the sampling distributions of the respective estimators generated from 5000 paired-sample bootstrap replications. These estimators are constructed by generating 5000 samples of 11 observations in which each observation in a sample is a common factor realization (from Exhibit 22) and a ratings class default rate pair chosen at random from the original sample (with replacement). For each of the 5000 samples, the respective estimators are calculated and the sampling distribution of the estimators is constructed.

Exhibit 23: PD Sampling Distribution Estimates for Selected Grades of Moody's Alpha-Numeric Rating Scale, 1998-2008

	sample estimate	sampling distribution characteristics for the rating grade PD estimate: mean	standard deviation	quantile: 99	95	90	10	5	1
Aa3 corrected	0.287	0.293	0.059	0.468	0.398	0.371	0.225	0.211	0.186
Aa3 simple	0.316	0.317	0.113	0.666	0.549	0.433	0.200	0.200	0.200
A1 corrected	0.275	0.279	0.048	0.411	0.364	0.344	0.222	0.209	0.186
A1 simple	0.272	0.272	0.069	0.487	0.415	0.344	0.200	0.200	0.200
A2 corrected	0.254	0.256	0.035	0.344	0.317	0.303	0.212	0.201	0.183
A2 simple	0.221	0.221	0.020	0.285	0.264	0.243	0.200	0.200	0.200
A3 corrected	0.255	0.256	0.033	0.342	0.313	0.300	0.215	0.206	0.189
A3simple	0.222	0.222	0.029	0.289	0.267	0.245	0.200	0.200	0.200
Baa1 corrected	0.307	0.309	0.042	0.420	0.385	0.366	0.259	0.249	0.231
Baa1 simple	0.318	0.318	0.089	0.594	0.503	0.422	0.216	0.213	0.206
Baa2 corrected	0.327	0.329	0.041	0.432	0.400	0.384	0.277	0.265	0.244
Baa2 simple	0.320	0.320	0.062	0.475	0.424	0.407	0.241	0.223	0.210
Baa3 corrected	0.380	0.385	0.068	0.568	0.508	0.477	0.303	0.283	0.254
Baa3 simple	0.447	0.447	0.143	0.826	0.710	0.644	0.268	0.239	0.219
Ba1 corrected	0.407	0.414	0.080	0.624	0.558	0.520	0.318	0.295	0.256
Ba1 simple	0.500	0.500	0.165	0.953	0.798	0.715	0.295	0.266	0.226
Ba2 corrected	0.529	0.536	0.090	0.773	0.689	0.651	0.427	0.397	0.342
Ba2 simple	0.600	0.599	0.129	0.910	0.818	0.768	0.432	0.399	0.315
Ba3 corrected	1.108	1.124	0.187	1.602	1.459	1.377	0.892	0.841	0.747
Ba3 simple	1.311	1.313	0.272	1.980	1.757	1.666	0.967	0.873	0.730
B1 corrected	1.098	1.128	0.254	1.808	1.583	1.468	0.819	0.753	0.634
B1 simple	1.489	1.489	0.337	2.255	2.039	1.921	1.056	0.925	0.716
B2 corrected	2.180	2.254	0.618	3.937	3.370	3.079	1.522	1.383	1.132
B2 simple	3.277	3.276	0.908	5.538	4.838	4.469	2.144	1.871	1.357
B3 corrected	4.421	4.518	1.030	7.198	6.301	5.865	3.233	2.902	2.410
B3 simple	5.915	5.916	1.468	9.602	8.489	7.852	4.063	3.656	2.897
Caa1 corrected	7.847	7.912	1.217	11.231	10.073	9.532	6.413	6.096	5.505
Caa1 simple	9.334	9.342	2.011	14.365	12.871	12.013	6.818	6.248	5.217
Caa2 corrected	15.620	15.719	1.847	20.363	18.853	18.128	13.400	12.760	11.720
Caa2 simple	17.435	17.443	2.837	23.889	22.146	21.074	13.803	12.889	10.719
Caa3 corrected	23.825	23.859	1.674	28.182	26.726	26.017	21.757	21.235	20.342
Caa3 simple	24.823	29.815	2.676	31.397	29.293	28.411	21.382	20.590	19.202

Corrected estimates use consistent estimates of the VAIRB common factor from Moody's 1920-2008 corporate ratings data (Exhibit 22) to construct the estimator given in expression (16). Sampling distributions are estimated using 5000 paired-sample bootstrap replications.

A comparison of the simple and corrected PD estimates and sampling distribution characteristics reported in Exhibit 23 reveals some interesting features. The first thing to notice is that for most credit grades, the corrected PD estimator is smaller than the simple PD estimator. This feature is intuitive as the common factor realization estimates include 7 of 11 years with below-average common factor realizations. Default rate observations are elevated from their unconditional PD rates and are adjusted downward by expression (16).

The corrected PD estimates are smaller than simple PD estimates for all grades except A1, A2, A3 and Baa2. This is easily explained. Grades A1, A2, A3 and Baa2 have only one observation in the 11-year sample with a non-zero default rate and the ZDROs are truncated to 20 basis points. The common factor correction adjusts observed default rates by adjusting an index that is transformed into a default rate through a standard normal CDF transformation. The PDF of the standard normal distribution for the index value associated with default rates below 20 basis points [$\Phi^{-1}(.002) = -2.878$], is very small. In contrast, the PDF values associated with index levels in excess of 20 basis points is not only larger, it is increasing as the default rate increases. The asymmetry of the normal distribution PDF in this region explains why the corrected PD estimates for grades A1, A2, A3 and Baa2 are larger than the uncorrected estimates. For 7 of the 11 sample observations, the common factor adjustment works to lower the index that determines the portfolio default rate, but in the region of $\Phi^{-1}(.002) = -2.878$, small declines in the index imply very little decline in the portfolio default rate. Conversely, for 4 of the 11 sample points, the common factor adjustment works to increase the index, and in the region of $\Phi^{-1}(.002) = -2.878$, the implied increase in portfolio default rates is large compared to the reductions generated by the corrections for the 7 above-average default rate years.

A second important feature of the corrected PD estimator can be seen comparing the confidence intervals (CIs) for the respective PD estimators calculated from the sampling distributions for each grade grade.[28] In all grades but two cases—the A2 and A3 grades, the width of the symmetric CI is smaller for the corrected PD estimate. For all other grades, the width of the CI for the corrected estimator is on average about 65 percent of the width of the CI associated with the simple PD estimator. The corrected PD estimators for grades A2 and A3 have wider CI intervals because of the data characteristics of these grades. Both grades have only one non-zero default rate observation, and the single non-zero default rate is very small in both cases (about 45 basis points). After ZDROs are truncated to 20 basis points, the simple PD estimator for these two credit grades has only a very small associated standard deviation.

XI. SUMMARY AND CONCLUSIONS

This paper has developed a new approach for estimating the parameters of the asymptotic Vasicek portfolio default rate model that is used to set minimum capital requirements for large, complex banking institutions under the Basel II AIRB framework. The approach produces consistent estimates of all VAIRB parameters using only time series data on a cross section of the failure rates from a consistent credit rating system. The method is applied to data on Moody's rated corporate credits over the period 1920–2008.

The results show that the VAIRB is incapable of explaining the variability of the default rates in Moody's data. From the perspective of capital allocation, an important finding is that the model significantly under-predicts default rates during "bust" phases of the credit cycle. This pattern of under-prediction is consistent with a missing transitory

[28] A confidence interval estimate is the difference between to two extreme quantile estimates, say the 99-percent and the 1-percent, or the 95-percent and the 5-percent.

common factor that affects default correlation only intermittently, similar to the frailty covariate in Duffie et al. (2009).

Bootstrap methods are used to construct estimates of the VAIRB model error distribution by credit grade. These distributions are used to estimate the capital needed as a buffer against model risk when the VAIRB is used to set capital requirements. The results show that capital needed as a buffer against model risk is substantial. Indeed, for lower-rated credits, model risk capital exceeds the capital needed to cover 99.9 percent of potential credit losses as estimated by VAIRB.

Estimates of the VAIRB common factor indicate that factor realizations are strongly autocorrelated. This finding implies that a long time series of default rates is needed to construct consistent estimates of the PD for a credit grade. The sampling distributions for VAIRB parameter estimates derived from small samples are constructed using the jackknife. The results show that small-sample estimators have high variability and that there is a substantial likelihood that small-sample PD estimates from will substantially underestimate the true probability of default. Consistent estimates of VAIRB common factors derived from long time series can be used to improve the accuracy of PD estimates (and confidence intervals) derived from small samples. Small-sample PD estimates that include a correction for the common factor realizations are constructed for credits rated using Moody's alpha-numeric scale over the period 1998-2008. For most credit grades, the correction for the common factor results in smaller PD estimates and narrower confidence intervals as compared to simple average default rate estimates over this sample period.

XII. APPENDIX: ESTIMATION OF UNCONDITIONAL DEFAULT RATES AND TIME EFFECTS WHEN DEFAULTS ARE AUTOCORRELATED

When the latent common factor is autocorrelated, expression (9) is isomorphic to the following specification:

$$\widetilde{Z}_{it} = a_i + \widetilde{u}_t + \widetilde{\upsilon}_{it} \tag{A1}$$

with

$$\widetilde{u}_t = \alpha\, u_{t-1} + \widetilde{\varsigma}_t\ , \qquad \alpha < 1 \tag{A2}$$

where *i* indexes the cross section of *N* groups and *t* indexes time. The mean-zero error components, $\widetilde{\upsilon}_{it}$ and $\widetilde{\varsigma}_t$, are independent and identically distributed over time. Consistent estimates of the unconditional default rates and latent factor realizations can be constructed from consistent estimates of a_i and realized values of $\widetilde{u}_t$ in expression A1.

Assuming a sample of observations on *T* time periods, repeated substitution yields

$$\begin{aligned}
\widetilde{Z}_{iT} &= a_i + \widetilde{\varsigma}_T + \alpha\,\widetilde{\varsigma}_{T-1} + \alpha^2\widetilde{\varsigma}_{T-2} + \ldots + \alpha^T\widetilde{\varsigma}_0 + \widetilde{\upsilon}_{iT} \\
\widetilde{Z}_{iT-1} &= a_i + \widetilde{\varsigma}_{T-1} + \alpha\,\widetilde{\varsigma}_{T-2} + \alpha^2\widetilde{\varsigma}_{T-3} + \ldots + \alpha^{T-1}\widetilde{\varsigma}_0 + \widetilde{\upsilon}_{iT-1} \\
\widetilde{Z}_{iT-2} &= a_i + \widetilde{\varsigma}_{T-2} + \alpha\,\widetilde{\varsigma}_{T-3} + \alpha^2\widetilde{\varsigma}_{T-4} + \ldots + \alpha^{T-2}\widetilde{\varsigma}_0 + \widetilde{\upsilon}_{iT-2} \\
&\ \ \vdots \\
\widetilde{Z}_{i1} &= a_i + \widetilde{\varsigma}_1 + \alpha\,\widetilde{\varsigma}_0 + \widetilde{\upsilon}_{i1}\,.
\end{aligned} \tag{A3}$$

The sample mean of $\widetilde{Z}_{it}$ is a consistent estimator of a_i, $\hat{a}_i = T^{-1}\sum_{t=1}^{T} Z_{it}$,

$$\lim_{T\to\infty} \quad \hat{a}_i \xrightarrow[a.s]{} a_i + \left(\frac{1}{1-\alpha}\right) E(\widetilde{\varsigma}_t) + E(\widetilde{\upsilon}_i) = a_i \tag{A4}$$

where *a.s.* almost surely indicates convergence.

If the sample size *T* is small and the autoregressive term α is large (i.e., close to 1), then the autoregressive error component will have a large effect on the sample mean estimator, and sample mean will be a poor estimator for a_i. In general, the larger the autoregressive parameter α, the longer the time series that will be needed to produce accurate estimates of the unconditional sample mean a_i.

One can construct consistent estimates of the time effects by using consistent estimates of each group sample mean. For a cross section of size *N*,

$$\tilde{Z}_{it} - \hat{a}_i = \tilde{u}_t + \tilde{\upsilon}_{it} \quad \text{for } i = 1, 2, 3, \ldots, N. \tag{A5}$$

The cross sectional average, $\hat{u}_t = N^{-1} \sum_{i=1}^{N} \left(\tilde{Z}_{it} - \hat{a}_i \right)$, is a consistent estimator of the realized value of the $\tilde{u}_t$ realization,

$$\lim_{N \to \infty} \hat{u}_t = u_t + \lim_{N \to \infty} \frac{\sum_{i=1}^{N} \tilde{\upsilon}_{it}}{N} \xrightarrow[a.s]{} u_t\,. \tag{A6}$$

The consistency of the $\hat{u}_t$ estimators is not altered by the imposition of the additional restriction that time effect estimates must sum to zero over the sample horizon.

In short time series applications, imposing the restriction $\sum_{t=1}^{T} b_t = 0$ will result in biased estimates of the time effects, for the actual sample will not be balanced between positive and negative deviations from the unconditional mean. As the length of the sample increases, the average effect of the common factor over the sample will more closely approximate zero, and the identifying restriction $\sum_{t=1}^{T} b_t = 0$ will be appropriate. The importance of partial cycles will diminish as *T* increases and the sample includes more complete cycles. At the limit, even if there are credit cycles in the default data, the restriction $\sum_{t=1}^{T} b_t = 0$ will hold exactly as $T \to \infty$.

REFERENCES

Agresti, A., and B. A. Coull (1998). “Approximate Is Better Than ‘Exact’ for Interval Estimation of Binomial Proportions.” *American Statistician* 52:119–26.

Altman, Edward and Duen Kao (1992). “The Implications of Corporate Bond Ratings Drift,” *Financial Analysts Journal*, 48:3, pp. 64-75.

Basel Committee on Banking Supervision (2006). *International Convergence of Capital Measurement and Capital Standards.* Bank for International Settlements.

Cameron, A. Colin, and Pravin K. Trivedia (2005). *Microeconometrics: Methods and Applications.* New York: Cambridge University Press.

Cantor, Richard, and Eric Falkenstein (2001). “Testing for Rating Consistency in Annual Default Rates.” *Journal of Fixed Income* (September): 36–51.

Cantor, Richard, David Hamilton, and Jennifer Tennant (2007). “Confidence Intervals for Corporate Default Rates.” SSRN Working Paper 995545.

Chernick, Michael R. (1999). *Bootstrap Methods, A Practitioner's Guide*. Wiley Series in Probability and Statistics.

Das, Sanjiv, Darrell Duffie, Nikunj Kapadia, and Leandro Saita (2007). “Common Failings: How Corporate Defaults Are Correlated,” *Journal of Finance* 62:93–117.

Das, Sanjiv, Laurence Freed, Gary Geng, and Nikunj Kapadia (2006). “Correlated Default Risk,” *Journal of Fixed Income* (Fall): 1–26.

Davison, A. C.; and D. Hinkley (2006). *Bootstrap Methods and Their Applications.* 8th ed. Cambridge U.K.: Cambridge Series in Statistical and Probabilistic Mathematics.

Diaconis, P., and B. Efron (1983). “Computer-Intensive Methods in Statistics”. *Scientific American* (May), pp.116–30.

Duffie, Darrell, Andreas Eckner, Guillaume Horel, and Leandro Saita (2009). “Frailty Correlated Default.” *Journal of Finance* 64 (5): 2089–2123.

Efron, B. (1979). “Bootstrap Methods: Another Look at the Jackknife.”. *Annals of Statistics* **7** (1): 1–26.

Financial Stability Institute (2006). “Implementation of the New Capital Adequacy Framework in Non–Basel Committee Member Countries.” Bank for International Settlements, Occasional Paper No. 6. http://www.bis.org/fsi/fsipaper06.pdf.

Finger, Chris (1999). “Conditional Approaches for CreditMetrics Portfolio Distributions.” *CreditMetrics Monitor,* pp. 14–33.

Gordy, Michael (2003). “A Risk-Factor Model Foundation for Ratings-Based Bank Capital Rules.” *Journal of Financial Intermediation* 12 (July), pp.199–232.

Güttler, Andre and Peter Raupach (2009). “The Impact of Downward Rating Momentum,” forthcoming, *The Journal of Financial Services Research.*

Hanson, Samuel, and Til Schuermann (2006). “Confidence Intervals for Probabilities of Default.” *Journal of Banking and Finance* 30:8, 2281–2301.

Kupiec, P. (2003). “Understanding the Expected Loss Debate.” *Risk* 11:11, pp. 29-32.

——— (2007a). “Financial Stability and Basel II.” *Annals of Finance* 3(1):107–30.

——— (2007b). “Capital Allocation for Portfolio Credit Risk.” *Journal of Financial Services Research* 32 (1–2): 103–122.

——— (2008a). “Basel II: A Case for Recalibration.” In *Handbook of Financial Intermediation and Banking,* edited by Anjan Thakor and Arnoud Boot. New York: North Holland.

——— (2008b). “A Generalized Single Common Factor Model of Portfolio Credit Risk.” *Journal of Derivatives* 15 (3): 25–40.

Lando, David, and Torben Skødeberg (2002). “Analyzing Rating Transistions and Rating Drift with Continuous Observations.” *Journal of Banking and Finance* 26:1-2, pp. 423–44.

Lopez, Jose (2004). “The Empirical Relationship between Average Asset Correlation, Firm Probability of Default, and Asset Size.” *Journal of Financial Intermediation* 13:265–83.

Moody’s Investor Service (2009). “Corporate Default and Recovery Rates, 1920–2008.” Moody’s Investor Service.

Pluto, Katja, and Dirk Tasche (2005). “Estimating Probabilities of Default for Low Default Portfolios,” memo, Deutsche Bundesbank.

Schönbucher, P. (2001). “Factor Models: Portfolio Credit Risks When Defaults Are Correlated.” *Journal of Risk Finance* 3 (1): 45–56.

Vasicek, O. (1987). “Probability of Loss on a Loan Portfolio.” KMV Working Paper. Published as “Loan Portfolio Value,” *Risk* (December 2003): 160–62.

Wald, A., and J. Wolfowitz (1940), "On a Test Whether Two Samples Are from the Same Population." *Ann. Math. Statist.* 11, pp.147–62.

Zeng, Bin, and Jing Zhang (2001). "An Empirical Assessment of Correlation Models." KMV Report (November 4).