Forecasting Default with the Merton Model (previously the KMV-Merton Model)

Sreedhar T Bharath
University of Michigan
Ross School of Business

Tyler Shumway
University of Michigan
Ross School of Business

May 16, 2006

Forecasting Default

- People have been forecasting default for decades
  - Classification: Altman (1968)
  - Structural: Merton (1974)
  - Hazard: Shumway (2001)
- KMV's implementation of Merton is interesting
- What does the KMV-Merton model contribute?

KMV and the Merton Model

- The KMV-Merton model is based on Merton's (1974) bond pricing model
- Developed by KMV corporation in the late 1980s
- Moody's bought KMV in 2002 for $210 million
- We call the model "KMV-Merton" because it is a nontrivial extension of Merton – credit KMV
- Others just call it a Merton model
- It is not exactly what MKMV sells - cheap version

The KMV-Merton Model

- The model uses market equity, equity volatility, and the face value of debt to infer the P(default)
- It recognizes that the market value of debt is unobservable – uses equity to infer debt value
- It is widely used in practice, new in academics
  - Vassalou and Xing (2004)
  - Duffie, Saita and Wang (2005)
  - Campbell, Hilscher and Szilagyi (2005)

Our Paper

- We ask how and why KMV-Merton works
- Hypothesis one: π_{Merton} is a sufficient statistic for predicting default
- Hypothesis two: We need π_{Merton} to construct a sufficient statistic for predicting default
- We also ask which parts of KMV calculation are important – can we cut corners?

Assessing KMV-Merton

- Our research strategy:
  1. Calculate KMV-Merton probabilities (π_{Merton})
  2. Propose a simple, "naive" alternative (π_{Naive})
  3. Estimate hazard models for time to default with π_{Merton}, π_{Naive} and other variables
  4. Look at the out of sample predictive power
  5. Estimate regressions for credit default swap (CDS) probabilities and bond yield spreads

The Merton Model

- Merton's assumptions:

  1. One zero-coupon bond with face value F and maturity T
  2. Firm value, V, geometric Brownian motion
  3. Other Black-Scholes-Merton assumptions

- Equity, E, is a call option on V with strike equal to F and maturity of T

Equity in the Merton Model

- Equity value is given by the Black-Scholes-Merton

$$E = V\mathcal{N}(d_1) - e^{-rT}F\mathcal{N}(d_2), \tag{1}$$

$$d_1 = \frac{\ln(V/F) + (r + 0.5\sigma_V^2)T}{\sigma_V T^{(1/2)}} \tag{2}$$

- Debt value is given by the value of a risk-free bond minus the value of a put written to equity

- Equity volatility follows

$$\sigma_E = \left(\frac{V}{E}\right)\frac{\partial E}{\partial V}\sigma_V = \left(\frac{V}{E}\right)\mathcal{N}(d_1)\sigma_V \tag{3}$$

Two Equations

- We have two nonlinear equations,

$$E = V\mathcal{N}(d_1) - e^{-rT}F\mathcal{N}(d_2),$$

$$\sigma_E = \left(\frac{V}{E}\right)\mathcal{N}(d_1)\sigma_V$$

- We can observe E, F, r, T; we can estimate σ_E

- We have two unknowns, or unobservables

- Solve the system for V and σ_V

- This is KMV's contribution - excellent idea

Distance to Default

- With V and σ_V, we can define the distance to default as a Z-score

$$DD = \frac{\ln(V/F) + (\mu - 0.5\sigma_V^2)T}{\sigma_V T^{(1/2)}}, \tag{4}$$

- Corresponding probability of default (or EDF) is

$$\pi_{\text{Merton}} = \mathcal{N}(-DD). \tag{5}$$

- This is not exactly what we (or KMV) actually do

Iterating on σ_V

- KMV claims that solving both equations simultaneously gives bad results
- Instead we (they) start with an initial σ_V
- Then we solve the BS equation for V given E each day for the previous year, using our σ_V
- We take the resulting time-series of V and calculate a new σ_V and μ
- We iterate in this way until σ_V converges

Naive Alternative

- We want to construct a measure that is similar to π_{Merton} without solving equations and iterating
- Somewhat arbitrarily, we define:
  - Naive $V = E + F$
  - Naive $\sigma_V = [E/(E+F)]\sigma_E + [F/(E+F)](0.05 + 0.25 * \sigma_E)$
  - Naive $\mu = r_{it-1}$
- We calculate naive DD similarly to DD, and we calculate the corresponding π_{Naive}
- This naive probability is simple to figure
- It captures the form and information of π_{Merton}

Hazard Models

- Having defined π_{Merton} and π_{Naive}, we need to determine which is a better forecaster

- We use the Cox proportional hazard model with time-varying covariates to assess the models

- The hazard rate is the conditional probability of failure at time t given survival until time t

Proportional Hazard Model

- The proportional hazard model stipulates that

$$\lambda(t) = \phi(t)[\exp(x(t)^{'}\beta)], \tag{6}$$

- We use standard techniques to estimate β

- Hazard models are probably the current state of the art for reduced form default models
  - Shumway (2001), Chava and Jarrow (2004)

MKMV vs Our Implementation

- We do not do everything exactly like KMV
  - Moody's KMV uses proprietary KV model
  - Uses historical data to define CDF of π_{Merton}
  - May use a proprietary formula to get F
- While we do not exactly match Moody's KMV, we match academic applications
- We can compare to some published KMV data

Table 2: Comparison with Moody's KMV

Correlation	Estimate
Rank Corr(Moody's EDF, Our π_{Merton})	0.788
Rank Corr(Moody's EDF, Our π_{Naive})	0.786
Rank Corr(Moody's σ_V, Our σ_V)	0.574
Rank Corr(Moody's σ_V, Our Naive σ_V)	0.853

- Based on 80 firms for which KMV EDF for August 2000 is published in *CFO Magazine*

Data

- We take all firms from CRSP/Compustat from 1980 through 2003
- We collect defaults from Altman and Moody's
- Sample has 1,016,552 firm-months, 1449 defaults
- Set $T = 1$, $F =$ debt in current liabilities $+$ $1/2$ of long-term debt
- Previous returns are from CRSP - past year
- Winsorize most variables at 1% and 99%

Table 3: Hazard Model Estimates

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
π_{Merton}	3.635***			1.697***	3.272***	0.230	
	(0.068)			(0.142)	(0.077)	(0.164)	
π_{Naive}		4.011***		2.472***		1.366***	1.526***
		(0.067)		(0.147)		(0.178)	(0.138)
ln(E)			-0.472***		-0.164***	-0.247***	-0.255***
			(0.014)		(0.015)	(0.024)	(0.023)
ln(F)						0.263***	0.269***
						(0.020)	(0.020)
$1/\sigma_E$						-0.506***	-0.518***
						(0.047)	(0.046)
$r_{it-1} - r_{mt-1}$						-0.819***	-0.834***
						(0.081)	(0.080)
NI/TA						-0.044***	-0.044***
						(0.002)	(0.002)

- π_{Merton}, π_{Naive} similar coefficients, significance
- Other significant variables – reject hypothesis one
- π_{Merton} not significant in Model 6 - reject two?

Table 4: Out of Sample Forecasts

Panel A: 1980 - 2003

350,662 firm-quarters, 1449 defaults

Decile	π_{Merton}	π_{Naive}	E	$r_{it-1} - r_{mt-1}$	NI/TA
1	64.9	65.8	35.7	44.4	46.8
2	15.1	14.3	17.5	25.1	23.8
3	6.0	6.7	14.3	9.2	10.6
4	4.6	4.1	9.1	5.4	5.9
5	2.9	2.4	6.1	2.9	4.2
6 - 10	6.5	6.7	17.3	13.0	8.7

- Sort firms each quarter into deciles based on probability and count defaults in each decile
- π_{Naive} performs as well as π_{Merton}
- Other sorts perform surprisingly well

Table 4: Out of Sample Forecasts

Panel B: 1991 - 2003

226,604 firm-quarters, 842 defaults

Decile	π_{Merton}	π_{Naive}	Model 6	Model 7
1	68.8	70.3	77.1	76.8
2	15.3	12.6	10.5	10.5
3	5.1	6.2	4.9	4.9
4	3.0	3.4	1.8	2.1
5	1.9	1.8	1.2	1.0
6 - 10	5.9	5.7	4.5	4.7

- Shorter horizon allows fitted Models 6 and 7
- Model with π_{Merton} performs slightly better
- We cannot completely reject hypothesis two

Table 5: Alternative Predictors

Panel C: Out of Sample Forecasts

Decile	$\pi^{\mu=r}_{\text{Merton}}$	$\pi^{simul}_{\text{Merton}}$	$\pi^{imp\sigma}_{\text{Merton}}$	π_{Merton}	π_{Naive}
1	60.0	65.1	84.1	80.7	83.0
2	17.7	15.0	8.0	9.1	9.1
3	8.0	7.7	4.6	3.4	5.7
4	4.1	3.4	0.0	5.7	1.1
5	3.4	3.2	1.1	0.0	0.0
6 - 10	6.8	5.6	2.2	1.1	1.1
Defaults	1,449	1,449	88	88	88
Firm-Quarters	350,662	350,662	36,274	36,274	36,274

- $\pi^{\mu=r}_{\text{Merton}}$ sets $\mu = r$, performs worse
- $\pi^{simul}_{\text{Merton}}$ avoids iteration, performs as well
- $\pi^{imp\sigma}_{\text{Merton}}$ uses option-implied σ_E, performs well

Table 6: CDS Spread Regressions

Dependent Variable: π_{CDS}

Variable	Model 1	Model 2	Model 3	Model 4
Const.	.03***	.03***	.03***	.16***
	(.0005)	(.0004)	(.0004)	(.007)
π_{Merton}	.05***		-0.001	.009**
	(.004)		(.003)	(.004)
π_{naive}		.13***	.13***	.06***
		(.007)	(.008)	(.008)
$\ln(E)$				-.007***
				(.0005)
$\ln(F)$				.003***
				(.0006)
$1/\sigma_E$				-0.014***
				(.0007)
r_{it-1}				-0.0012
				(.001)
Obs.	3833	3833	3833	3833
R^2	0.10	0.26	0.26	0.40

- Including π_{Naive} makes π_{Merton} lose significance

Table 7: Bond Yield Spread Regressions

Dependent Variable: Bond Yield Spread

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Const.	125.77***	118.27***	126.2***	153.11***	147.28***	153.8***
	(16.55)	(16.59)	(16.59)	(17.43)	(17.36)	(17.46)
Omitted regressors – See the table in the paper						
π_{Naive}	.5***		.49***	.63***		.6***
	(.03)		(.03)	(.04)		(.04)
π_{Merton}		.08***	.007		.19***	.1***
		(.008)	(.007)		(.02)	(.01)
Rating Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Industry Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	61776	61776	61776	51831	51831	51831
R^2	0.70	0.70	0.70	0.72	0.71	0.72

- Including π_{Naive} makes π_{Merton} lose significance

Conclusion

- π_{Merton} is a decent default predictor, however
- We reject hypothesis one – π_{Merton} is not sufficient
- We can almost reject hypothesis two – π_{Merton} is useful with other predictors
- While the iterative procedure is not useful, the Z-score functional form is fairly useful
- It would be great to have MKMV data to perform similar tests - too expensive for us