The Journal of Behavioral Finance Volume 4, Issue 1, Spring 2014 62-77

ISSN: 1551-9570

How Expert Judgment Reacts to a Major Financial Crisis: An Analysis's of the 1925-1933 Bond Ratings of John Moody

John T. Donnellan
New Jersey City University

Berry K. Wilson
Pace University

Abstract

This study analyses Moody's railroad bond ratings over the period 1925-1933 to examine the stability of Moody's ratings during this critical period in U.S. economic history. One core principle applied by Moody was "results of the decade," [ROD] under which Moody derived quantitative measures from a 10-year moving average of the underlying credit data. Although related to "rating through the cycle [TTC]," ROD was a more explicit way of capturing secular trends in bond quality.

Moody decomposed the ratings construction process into two components. The first constructed a statistical rating from two risk factors. These were: (1) *security*, a measure of a bond's solvency, and (2) *salability*, a measure of a bond's liquidity. Moody then combined these two factors with his personal judgment to construct the final rating. Our study analyzes the stability of both the statistical rating and Moody's judgment. Our results contradict the ROD principle, and show that Moody made substantial changes in the rating process during the Great Depression.

"**The Results of the Decade**: In every figure presented, the yearly average of every factor is presented, for each year of the decade, and then these yearly averages are themselves averaged, thus showing *the average results for the entire decade*. By this method, a basis is reached which strongly emphasizes the element of **permanency and stability**, …" [emphasis added (Moody 1909)]

Conservative judgment: "The position of the issue is the results of the decade, thereby enabling one to arrive at a very conservative judgment of the strength and value of the given security".[emphasis added (Moody 1913)]

Introduction

John Moody single-handedly founded the security-rating industry in 1909. The industry quickly became influential. One significant segment of Moody's business was the large number of banks throughout the U.S., which held railroad bonds as (interest-bearing) secondary reserves. As such, bankers were concerned with both the solvency and liquidity of the railroad bond issues they invested in. However, given the multitude of American railroad bonds, and the small size and multitude of most U.S. banks, bankers individually lacked the resources and incentives to pursue their own in-xhouse security analysis. Thus the great demand for rating services in the U.S. (Wilson 2011).[1] This paper is in contrast to Wilson (2011) since it utilizes statistical residuals to explore the expert judgment aspect that John Moody incorporated into his bond ratings.

Moody based construction of his bond ratings explicitly on two factors which he termed *security* and *salability*. The *security* factor was measured using an interest-coverage ratio and reflected concern over the solvency of the issue. The *salability* factor reflected the marketability or liquidity of the issue. Thus these two factors satisfied the needs of bankers in selecting appropriate secondary reserves. Moody combined these two factors into a *statistical rating* that was then the quantitative basis of the rating.

In addition to the statistical rating, judgment also played a major role. Moody's explained:

> "The percentages showing the factors of safety, etc. serve as a general guide, but the rating given is, in many cases, affected by other considerations not shown in figures, such as character of management and of traffic, general position of the railroad system, policy of the company in maintenance and other expenses, and in other ways." (Moody 1920)

This research examines Moody's statistical ratings process and judgment over the years 1925-1933. We utilize an ordinal regression approach to capture the statistical ratings model used by Moody.

To examine the impact of Moody's judgment, our study analyzes the residuals from our estimated statistical ratings model. As suggested by the quote above, Moody's judgment likely reflected his experience with a large number of factors and circumstances that affected the quality of railroad bonds. If there were a multitude of such factors each with small effect, then we would expect the cross-section of Moody's judgment across railroad bonds to be approximately normally distributed. That is, the security and salability factors were the major influences behind bond quality, while Moody's judgment reflected the impact of a multitude of small factors and

[1] Wilson, B. (2011) finds that railroad securities were considered high quality at the time and were the preferred securities as collateral in the call loan market.

circumstances. In contrast, if Moody's judgment reflected only a small number of significant factors, then the distribution of his judgment might show significant skewness or kurtosis. Skewness in the distribution of Moody's judgment might reflect the anticipated differential impact of economic shocks on low-rated versus high-rated railroad bonds. For example, the shocks of the Great Depression might have been expected to have greater impact on low-quality railroads, as opposed to high-quality railroads, i.e., those railroad lines with higher-quality, more-dedicated traffic. The result would be a minority of high-rated bonds, plus a large number of lower-rated bonds reflecting the impact of the credit shocks of the 1930s. In contrast, kurtosis might reflect Moody's anticipation of the impact of extreme (i.e., 3-sigma, 4-sigma, 5-sigma, etc.) shocks on ratings.

Examining estimates of Moody's judgment can yield insights into whether Moody anticipated the extreme economic shocks of the Great Depression, and how Moody's judgment reflected ongoing events. The analysis will focus particularly on the year-by-year changes in the distribution of Moody's judgment, which will help control for fixed credits effects, and help to isolate the impact of changing credit conditions on Moody's judgment. This approach should yield the most useful insights into the judgment process.

Beyond examining construction of the statistical rating and Moody's judgment, the analysis tests the stability of these ratings components over the study period. One core principle applied by Moody was ROD under which Moody derived quantitative measures from a 10-year moving average of the underlying credit data. For the ROD principle to hold, the rating model coefficients would have to remain roughly constant. The study period 1925-33 provides a strong test of the stability hypothesis. The analysis also examines the stability of Moody's judgment. To repeat, a ROD approach would suggest that Moody's judgment reflected secular rather than cyclical changes in bond quality.

An analysis of Moody's ratings from the Great Depression era provides an important context for analyzing the 2008 subprime crisis, which has brought increased scrutiny to how to "fix" the bond ratings of the U.S. bond rating agencies. U.S. rating agencies were instrumental in underwriting the subprime mortgage-backed securities market leading up to the subprime crisis. Some problems with ratings include (1) the conflict of interest inherent in the "issuer pay" model used by rating agencies, (2) the ratings inflation that has occurred as rating agencies compete for ratings business, and (3) the lack of transparency on what credit ratings actually measure and what purpose they are intended to serve. Cantor and Packer (1994) present evidence that ratings inflation has occurred at least since the early 1970s. Wolfson and Crawford (2010) cite that "the United States Senate has accused the credit rating agencies of flawed methodology, weak oversight by regulators, conflicts of interest and a total lack of transparency." Relatedly, the Justice Department sued Standard & Poor's Ratings Services (S&P), alleging that S&P ignored its own standards to rate mortgage-backed securities.[2]

Our analysis shows that Moody's judgment was highly pro-cyclical during the 1920-30s. During the 1920s his judgment tended to "inflate" railroad bond ratings. Then, during the 1930s his judgment tended to downgrade securities below what the statistical ratings would suggest. This

[2] "U.S. Sues S&P Over Ratings; Justice Department Says Endorsements of Risky Mortgage Bonds Fueled Crisis." By J. Eaglesham, J. Neumann and E. Perez, *Wall Street Journal*, Feb. 5, 2013.

result appears to parallel the high ratings that U.S. ratings firms gave subprime debt leading up to the 2008 subprime crisis, and then the rapid downgrades that followed the crisis.

Also, our regression results show model R-squares ranging from 86% to 92.2%, indicating that the ratings were highly model driven. Yet Moody's statistical model was relatively simple, based on a bond liquidity measure and an interest coverage ratio. With the more recent subprime ratings, little analytical work went into establishing subprime ratings, relying instead on the tranching of mortgage assets into risk classes and the hope that diversification would provide some protection to senior tranches. These examples seem to indicate that the technology of ratings has not advanced much since the Great Depression.

Finally, the paper is divided into the following sections. Section I reviews relevant past literature. Section II describes the ratings construction process followed by Moody and Section III describes the data set and the hypotheses. Section IV discusses our ordinal regression approach then presents our cross-sectional regression results then presents panel (cross-sectional-time series) results. Our conclusions follow in Section V.

I. Literature Review

Rating TTC has been given various definitions by different researchers. The present study likely is the closest to Moody's original definition. As discussed above, "results for the decade" meant that credit-risk measurements, and the *security* measure in particular, were based upon a 10-year average of the underlying data. This approach is illustrated and discussed in the next section. Moody is less clear on how the judgment he impounded into ratings was based on a 10-year perspective. As well, this study defines rating stability in terms of the constancy of the estimated statistical ratings model and judgment components of Moody's ratings.

Amato and Furfine (2003) describes the rating process as a "threshold model with overshooting". Their results suggest that most often agencies simply monitor without rating changes. However, when credit conditions do significantly change, the rating agency "overreacts," perhaps reflecting excessive optimism or pessimism during economic upturns and downturns, respectively. The result is that when ratings change, they tend to be excessively procyclical.

Altman and Rijken (2005)study ratings stability in terms of the slow adjustment of ratings to new information. The authors find that over 1981-2001, credit quality can be decomposed into permanent and temporary components, and that rating agencies focus exclusively on the permanent component. The study also reports that the TTC method delays rating downgrades by 0.56 years and upgrades by 0.79 years. The TTC method results in an 8% drop in the accuracy of predicting one-year defaults.

Loeffler (2006) creates a separation between permanent and transitory components of credit risk, and shows that agency ratings focus on the permanent component. This result may explain the slowness of ratings to react to credit news and the low short-term forecast accuracy of ratings. Loeffler (2004) argues that this ability to identify the permanent component of credit risk seems to be a core competence of rating agencies. In contrast, Kiff, Kisser and Schumacher (2013) emphasize that the TTC approach is prone to "procyclical rating cliff effects" and is inferior in predicting defaults compared to point-in-time approaches.

These studies indicate that rating TTC has taken differing definitions in different studies. The present study uses Moody's original concept where ratings were based on a 10-year moving average of credit risk, which Moody termed ROD. This approach was in effect during the 1920-30s. Below, we decompose Moody's ratings into statistical rating and judgment components. The study then defines ratings stability as the insensitivity of these two components to credit-cycle movements, in the sense that a 10-year moving average should show minimal cyclical effects.

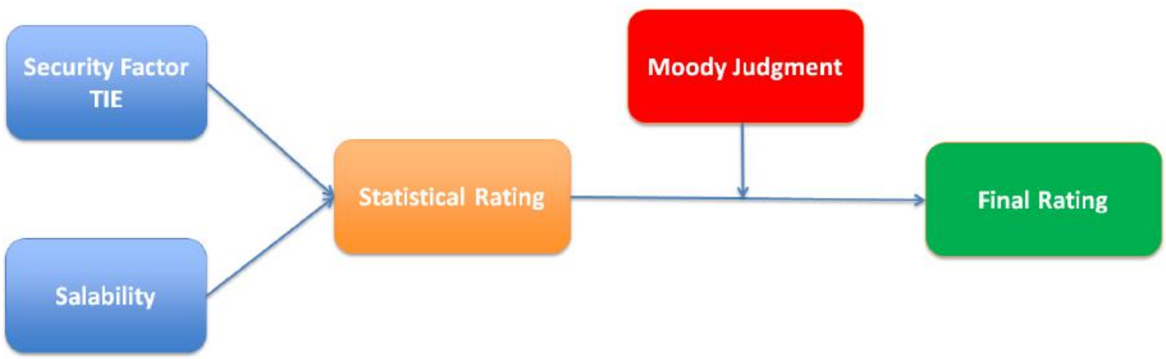
II. Construction of Moody's Bond Ratings

Moody's annual volumes described a two-part procedure that Moody followed to determine bond ratings.[3] This process is illustrated in Figure I.

Figure I diagrams the procedure used by Moody in establishing his bond ratings. Two risk factors: Security or TIE (Times Interest Earned) and salability were used as the basis for establishing a bond's "Statistical Rating." The security factor measures the railroad line's default risk, while the salability factor measured the security's liquidity risk.

To the statistical rating Moody added his own judgment about the bond's quality to yield the bond's final rating. Our analysis uses the two risk factors, whose values were recorded in the annual ratings manuals, to reconstruct and analyze Moody's judgment.

Figure I: **Construction of Moody's Ratings**



To reiterate, Moody first constructed a "statistical rating" that was explicitly based on two factors, which Moody termed *security* and *salability*. The *security* factor reflected the *earnings power* of the railroad line and was measured as a type of interest-coverage ratio. In contrast, the *salability* factor was a measure that reflected the security's liquidity. The security factor, termed Times Interest Earned (TIE), is calculated as follows.

TIE = AIA/CIR

Where AIA is average income available and CIR is the cumulative income required meeting the interest obligations of the security and all securities ranked higher.[4]

Figures II and III provide a study summary of these two risk factors that were the focus of Moody's statistical rating process.

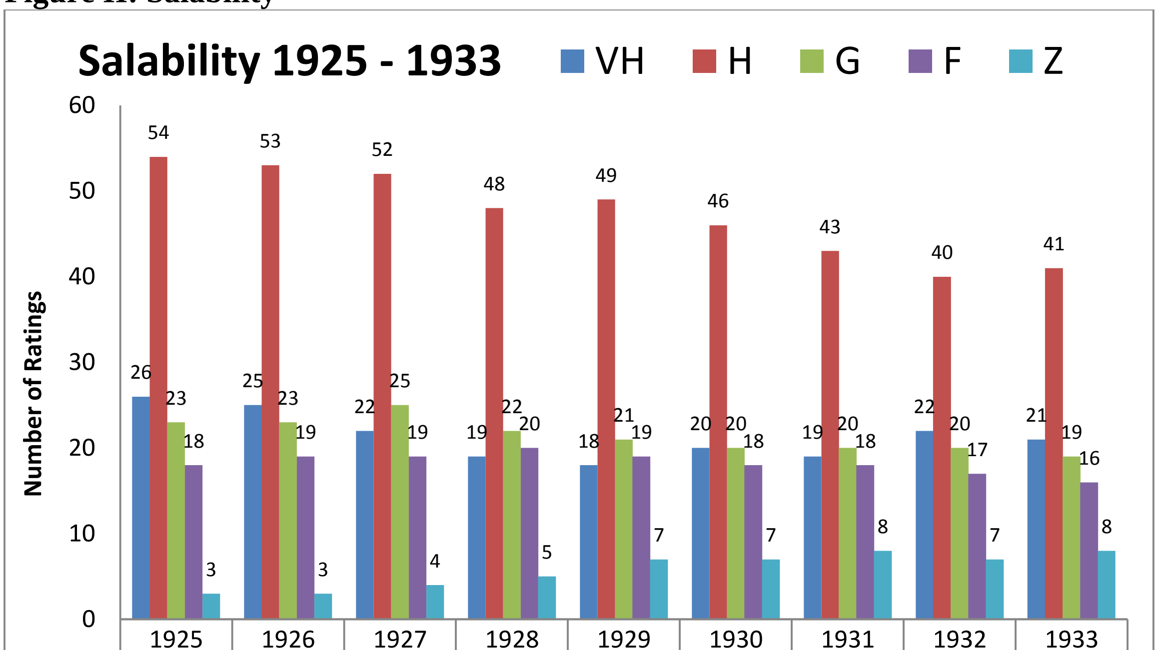
[3] Moody's bond ratings date to 1909, the first annual publication date of Moody's *Analysis of Railroad Investments* (Wilson, 2011). Moody's rating system initially included 14 rating categories: Aaa, Aa, A, Baa, Ba, B, Caa, Ca, C, Daa, Da, D , E and F. Ratings E and F were dropped in 1923. Ratings Daa, Da and D were dropped in 1930 (Fons 2004).

[4] Prior to the 1927 manual, Moody used the "Factor of Safety" as his measure of the *security* factor. It was calculated as Factor of Safety = 1 - 1/TIE= AIA/CIR

These two factors: the bond's security and salability were combined to yield what Moody termed the security's "statistical rating."[5] Therefore, the statistical rating represented the analytical basis for the bond ratings. Although Moody was mute on the exact process of how these two variables were combined, the present study adopts a flexible probabilistic model to capture this phase of the rating process.

Next, Moody combined his personal judgment concerning the quality of a particular security in combination with the "statistical rating" to determine the final rating of the security. As emphasized by Moody[6]: "The position of the issue is the ROD, thereby enabling one to arrive at a very conservative judgment of the strength and value of the given security". (Moody 1913)

Figure II: Salability[7]



	1925	1926	1927	1928	1929	1930	1931	1932	1933
VH	26	25	22	19	18	20	19	22	21
H	54	53	52	48	49	46	43	40	41
G	23	23	25	22	21	20	20	20	19
F	18	19	19	20	19	18	18	17	16
Z	3	3	4	5	7	7	8	7	8

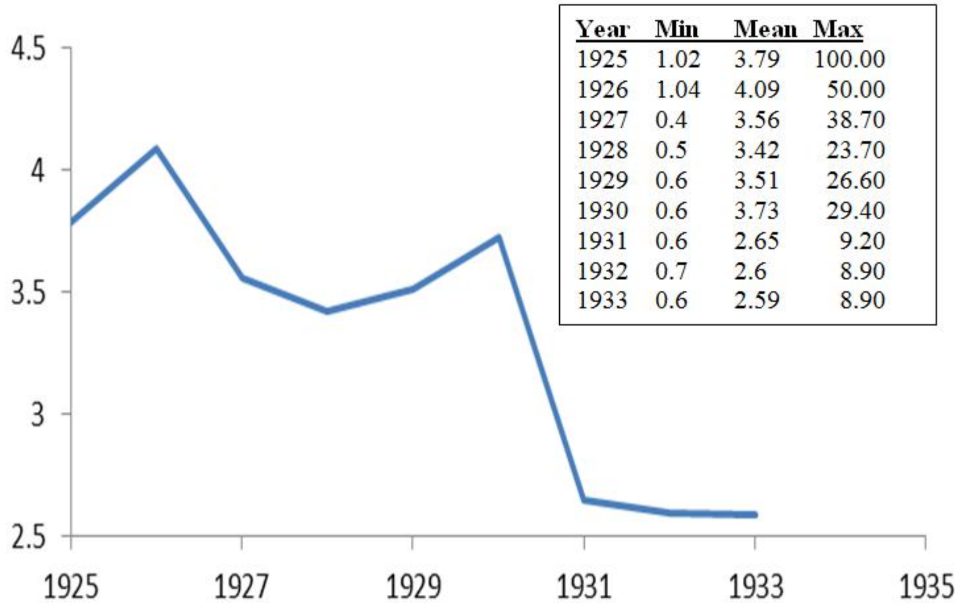
Legend: VH=Very High; H=High; G=Good; F =Fair; Z=Other [Poor, Low, Inactive]

[5] By Moody's definitions, (Aaa, Aa and A)-rated securities were little affected by the railroad's earnings-power risk, while (Baa, Ba and B)-rated securities were somewhat affected, and (Caa, Ca and C) or lower rated securities were largely affected (and thus "speculative") by the railroad's earnings-power risk.

[6] Moody did make some changes in the rating approach over the study period: 1925-1933. As mentioned above, in the 1927 ratings manual Moody changed from using the ssecurity factor: "Factor of Safety" to using: "Times Interest Earned". In 1934 Moody changed his measure of salability to the annual volume of NYSE bond market trades in the particular security. Before 1934 Moody constructed his liquidity measure from conversations with security dealers.

[7] Source: Moody's analyses of investments: Steam railroads 1925 through 1933

Figure III: Times Interest Earned (TIE) 1925 to 1933



Year	Min	Mean	Max
1925	1.02	3.79	100.00
1926	1.04	4.09	50.00
1927	0.4	3.56	38.70
1928	0.5	3.42	23.70
1929	0.6	3.51	26.60
1930	0.6	3.73	29.40
1931	0.6	2.65	9.20
1932	0.7	2.6	8.90
1933	0.6	2.59	8.90

Calculation: TIE = AIA/CIR, where AIA is average income available and CIR is cumulative interest required. Most of the decline in TIE resulted from the decline in average income available (AIA) over the period, i.e., the declining profitability of these railroad lines.
Data source: Moody's analyses of investments: Steam railroads 1925 through 1933.

III. Data Sampling and Hypothesis

All study data is taken from annual volumes 1925-1933 of *Moody's Analysis of Investments - Steam Railroads.* In particular, the study uses *Table D: Bond Record and Rating*, which displays the solvency and liquidity data that served as the basis for the statistical rating. The construction of a typical "Table D" is illustrated in Table I.

Table I – Moody Manual Section Table D

An example of the data table D presented in the Moody's annual manual for *Atchison, Topeka & Santa Fe Railroad*. This particular table is for the AT&SF line from the 1925 manual. The columns: (1) describes the particular bond issue, (2) annual payment months, (3) maturity date, (4) level of the lien (1st, 2nd, general, not mortgaged) and number of miles of the railroad line that backs the lien, (5) the dollar value of the security outstanding per mile of property, (6) average income available (AIA) to pay (7) interest required. A 10-year average income is used. Factor of safety is then calculated as (AIA-CIR)/AIA. Column (8) gives Salability and (9) the Rating that results from the security's Security and Salability measures. Note that construction of the table assumes that required interest is paid from security (1) to security (13), representing the security's priority of its claim. Thus, the CIR in calculating the Factor of Safety is the cumulative interest required as one moves from the top to bottom of the table.

TABLE D.—Bond Record and Ratings (Based on 10-Year Results, Per Mile of Road)

Name of Issue	Interest Payable	Maturity	Lien on Miles	Outstanding per Mile	Average Income Available per Mile	Interest Required per Mile of System	Factor of Safety	Security	Salability	Rating
1. Chic.,Santa Fe&Cal. 1st 5s	J&J	Ja 1937	(1st) 439	$1,276		$3	78%	Main	High	Aaa
2. A.T.&S.F.East. Okl.	M&S	Mr 1928	(1st) 476	20,174		42	78%	"	Very high	Aaa
3. A.T.&S.F.Trans.S.Line 4s.	J&J	Jl 1958	(1st)1113	20,256		99	78%	"	"	Aaa
4. Hutch. & So. Ry. 1st 5s. . .	J&J	Ja 1928	(1st) 141	1,362		1	78%	Branch	High	Aaa
5. San Fran.&San Jo.V.1st 5s	A&O	O 1940	(1st) 375	15,637		32	78%	Main	"	Aaa
6. Santa Fe,Prescott&P.1st5s	M&S	S 1942	(1st) 193	25,596		27	78%	Branch	"	Aaa
7. Prescott & Eastern 1st 5s. .	A&O	Ap 1928	(1st) 26	8,615	$5,740	1	78%	"	"	Aaa
8. A.T.&S.F. general 4s.	A&O	O 1995	(gen) 8539	17,641		664	78%	General	Very high	Aaa
9. A.T.&S.F. adjustment 4s. .	M&N	Jl 1995	(gen) 8539	6,013		226	78%	"	"	Aaa
10. A.T.&S.F.Cal.Ariz.L.4½s. .	M&S	Mr 1962	(gen) 749	24,730		92	78%	"	"	Aaa
11. A.T.&S.F. Rocky Mt. 4s. .	J&J	Ja 1965	(1st) 101	29,703		13	78%	Branch	High	Aaa
12. A.T.&S.F. convertible 4s. .	J&D	Je 1955	Not mtg.	......		35	78%	Debent.	Very high	Aaa
13. A.T.&S.F. convertible 4s. .	J&D	Je 1960	"	......		2	78%	"	"	Aaa

For the study's data collection purposes, the bond listed at the bottom of a particular Table D was selected for analysis, since it represented the lowest ranking security backed by the financial results for that particular railroad line. This bond then reflected the highest credit risk among the railroad line's bonds listed in the table.[8] In subsequent years, if this bond was dropped from the table, a replacement bond was chosen from the bottom of that year's table D for the railroad line. Over the nine study years: 1925-33, 19 bonds were dropped from the analysis due to (1) discontinuance of Table D (15 such bonds), or (2) bonds retiring without a replacement (4 bond).

Our data selection procedure selected one bond per railroad line to focus our analysis on Moody's judgment in ratings. Since all bonds backed by a common railroad line would reflect the same judgment concerning t (White 2010), analyzing one such bond is sufficient to capture Moody's judgment effects. Our sampling procedure resulted in a total of 124 bonds (drawn from 124 railroad lines), essentially all of the Table D's in the manuals.

The period 1925-1933 was chosen to provide a strong test of the stability of Moody's ratings during a period of prosperity and subsequent deep depression. If Moody's bond ratings were set by ROD (alternatively TTC), then we would expect to find only slow changes in ratings over time. Moody's approach to ratings was discussed above. By ratings stability we hypothesize that the parameters of the statistical ratings model remains constant over the study period, and that Moody's judgment likewise remained neutral to changes in the business cycle. Although considered to have expert judgment in railroad systems (Rodgers 2011), Moody could not possibly have anticipated how "bad" the Great Depression would become. Expert judgment in finance and economics is not always accurate. James Shanteau (1992) conducted a study on financial decision making and concludes "Experts are inadequate decision makers. That has been reinforced in past studies (Loeffler 2004) which have reported deficiencies in calibration (subjective-objective comparability) and coherence (internal consistency) of probability judgments. Furthermore, experts are apparently unaware of these various shortcomings." Today, Moody Analytics uses 'expert judgment' modeling in decision making that uses sophisticated algorithms not available to John Moody during the Great Depression.

As discussed above, Moody based his statistical rating on the security and salability factors. Our first hypothesis is that the parameters of the statistical ratings model remained constant

[8] In a few cases the sample data was incomplete for the bottom-most bond, then data was taken for the next-lowest bond in the table with complete data.

throughout the study period. That is, if Moody did indeed rate securities by the "ROD," then the statistical ratings model should show only slow changes through the business cycle.

In addition to changes in the statistical model, the economic fluctuations of the study period may have induced Moody to change how judgment was incorporated into the ratings. Moody had lived through significant economic recessions, including the recession of 1893 and the panic of 1907. Our study results should reveal how his judgment changed in response to the Great Depression.

IV. Approach and Results

An ordinal regression approach is adopted to account for the ordinal scale of the bond ratings.[9] Ordinal regression is a flexible approach to estimating the structural ratings model used by Moody. In the regression analysis, the *security* factor is incorporated as a continuous variable, and the *salability* factor is incorporated as a categorical regression variable.

The ordinal regression analysis first investigates the extent to which the two explicit factors: security and salability explain Moody's assignment of ratings. These two factors combine to create the security's "statistical rating." In particular, the analysis is interested in the stability of the estimated statistical rating model over the 1925-33 study periods. By ratings stability we essentially hypothesize that the parameters of the structural model remain constant over the cycle, which includes the rapid economic progress of the 1920s and the depth of the depression during the early 1930s.

Moody's bond ratings also incorporated Moody's personal judgment. To the extent that the ordinal regression analysis captures the statistical rating, the residuals of the ordinal regression analysis will provide estimates of the judgment that Moody incorporated into the ratings. Analyzing these residuals cross-sectionally over 1925-1933 will allow the analysis to assess: (1) the extent that judgment influenced the ratings and (2) how Moody's judgment responded to the depression years.

This section presents cross-sectional regression results over the sample period 1925-1933. A cross-sectional analysis allows the study to assess how the model coefficient estimates changed year-by-year through the 1920s and the depths of the 1930s. ROD implies that the statistical ratings model and its parameters should show minimal changes over the business cycle. Likewise, Moody's judgment should show minimal changes TTC. However, the Great Depression of the 1930s might have been largely unprecedented in Moody's lifetime experience.[10]

The results of the analysis are given in Table II. Results are given separately for each cross-section of data. That is, separate analyses are presented for each year over 1925-1933, respectively. First, the threshold parameter estimates represent the underlying frequencies of the ratings categories in estimating the probability model. Then, TIE factor adjusts the raw frequencies for the impact of the TIE variable. The TIE coefficient estimates of (-0.14, -.08, -0.41, -0.24, -0.28, -0.12, -0.53, -0.49, -0.51), respectively, indicate that there is an overall direct relationship between TIE and bond ratings. That is, a higher TIE value is associated with higher ratings. Note that the coefficient estimates appear to become more negative over the final three years. In the next section this shift will be formally tested.

[9] The PLUM procedure of SPSS was used to perform the ordinal regression.

[10] Friedman and Schwartz (1973) note that the Panic of 1907 was a short but severe financial crisis and the Panic of 1893 lacked the severe unemployment of the Great Depression.

Table II: Cross-Sectional Ordinal Regression Results

	1925	1926	1927	1928	1929	1930	1931	1932	1933
Threshold:									
R=1	-3.68***	-1.02	-2.09*	-1.47	-2.84***	-1.63*	-4.47***	-3.99***	-4.19***
R=2	-2.99**	-0.48	-1.24	-0.72	-2.06**	-1.13	-3.21***	-3.01***	-3.65***
R=3	-2.40*	0.03	-0.47	-0.03	-1.41*	-0.64	-2.71***	-2.49***	-3.04***
R=4	-1.87	0.53	0.10	0.73	-0.81	-0.07	-1.89**	-1.62**	-2.06***
R=5	-1.36	1.06	0.83	1.37	-0.11	0.46	-1.15	-0.66	-1.11*
R=6	-0.54				0.63	1.09	-0.10	0.37	-0.26
Location:									
TIE	-0.14**	-0.08*	-0.41***	-0.24***	-0.28***	-0.12**	-0.53***	-0.49***	-0.51***
Sal_F	-1.58	0.64	1.12	1.19	-0.06	0.18	-0.68	-0.26	-0.37
Sal_G	-2.06*	0.15	0.56	0.66	-0.38	0.11	-0.62	0.11	0.11
Sal_H	-2.40*	-0.35	0.08	-0.19	-0.97	-0.89	-1.20*	-0.34	-0.67
SalVH	-2.94**	-0.78	-0.51	-0.39	-1.18	-1.00	-1.65**	-0.71	-0.73
N	119	109	114	107	108	104	96	95	94
R^2	90.6%	77.7%	95.2%	94.1%	95.8%	65.1%	96.4%	96.0%	90.9%

Model Specification

Y is the ordinal response variable with J=1 (Aaa), =2 (Aa), …, to J=7 (Caa, Ca, and C). The model assumes that values of Y are drawn independently from a multinomial distribution. Let $\gamma_{ij} = \text{Prob}(Y \leq j|x_i)$, j=1 to J, be the cumulative probability of observation i being drawn from rating categories less than or equal to j, where x_i is a vector of independent variables. Note that $\gamma_{iJ} = 1$, hence only the first (J-1) γ's are needed in the model. Then γ_{ij} is given as: $\gamma_{ij} = \text{inv link}[(\theta_j - \boldsymbol{\beta}^T\mathbf{x_i}) / \sigma]$

where "inv link" is the inverse function of the link function, $\boldsymbol{\theta}$ is a vector of threshold parameters, $\boldsymbol{\beta}^T\mathbf{x_i}$ is the "regression" part of the model, and σ is a scaling parameter. In particular, a complementary log-log link function was used as the link function, and the SPSS PLUM procedure was used to perform the ordinal regression. We report the Cox and Snell Pseudo R^2 measure here.

Legend: R1=Aaa; R2=Aa, R3=A; R4=Baa; R5=Ba, R6=B; R7=Other (all C ratings), Sal_F =Fair; Sal_G=Good; Sal_H=High; Sal_VH=Very High; Sal_Z=Other (Poor, Low, Inactive).

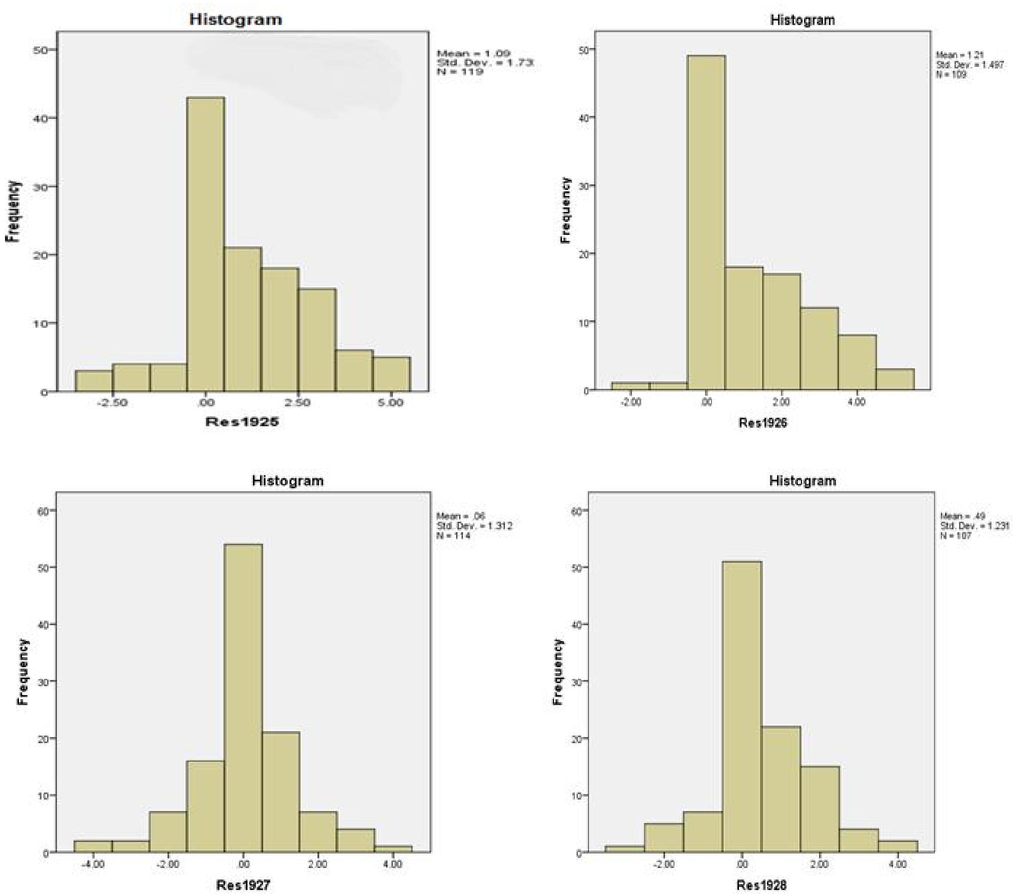
In the analysis the salability variable is discrete with levels: (F fair, G good, H high, VH very high, Z other), which are represented as dummy variables: SalF, SalG, SalH, and SalVH. SalZ is incorporated into the baseline case. Overall, the coefficients for SalF, SalG, SalH and SalVH show a trend of increasing from the baseline year of 1925. This conclusion will be formally tested in the next section.

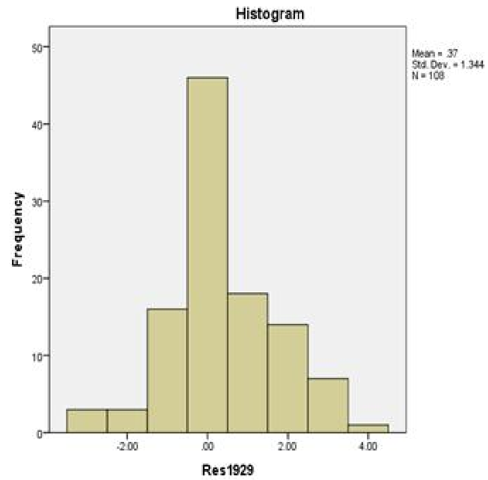
Summary statistics concerning the first four moments are illustrated in Table III and the cross-sections of residuals are diagrammed in Figure IV. Table III illustrated results indicate that (1) the residual volatility did not meaningfully increase during the depression years; (2) the skewness of the residuals did not show a pattern significantly different from zero over the data sample period, and (3) the kurtosis of the residuals did not show a significantly non–normal pattern over the sample period. We interpret these results as follows. First, the means indicate a down shift in ratings TTC, and in particular during 1931-33, even after controlling for statistical-ratings model changes in the ratings. Therefore, Moody's judgment appears to have accelerated the downgrade cycle in response to the Great Depression years.

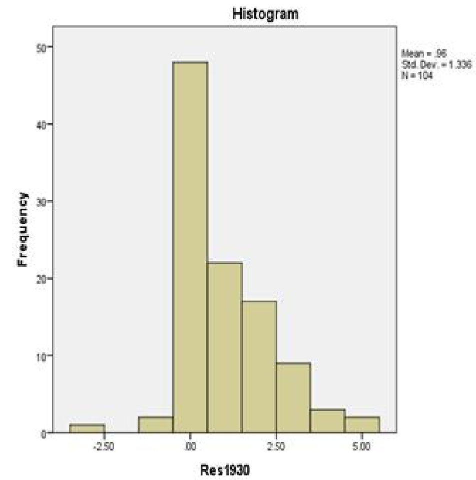
Table III: Summary Statistics Residual Analysis from the Ordinal Regression Results

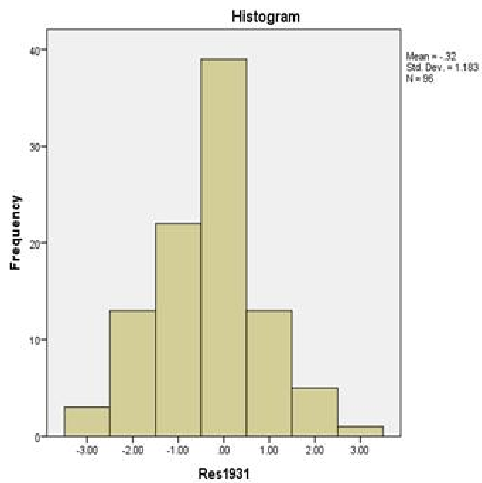
	1925	1926	1927	1928	1929	1930	1931	1932	1933
Mean	1.092	1.211	.0614	0.486	0.370	0.962	-.323	-0.463	-.3936
Std. Err.	(0.159)	(0.143)	(0.123)	(0.119)	(0.129)	(0.131)	(0.121)	(0.130)	(0.139)
S.D.	1.732	1.497	1.312	1.231	1.344	1.336	1.183	1.270	1.346
Skewness	0.243	0.794	-0.163	0.295	0.164	0.794	0.033	-0.594	-0.164
	(0.222)	(0.231)	(0.226)	(0.234)	(0.233)	(0.237)	(0.246)	(0.247)	(0.249)
Kurtosis	0.044	-0.228	1.705	0.911	0.398	1.045	0.160	0.063	0.393
	(0.440)	(0.459)	(0.449)	(0.463)	(0.461)	(0.469)	(0.488)	(0.490)	(0.493)

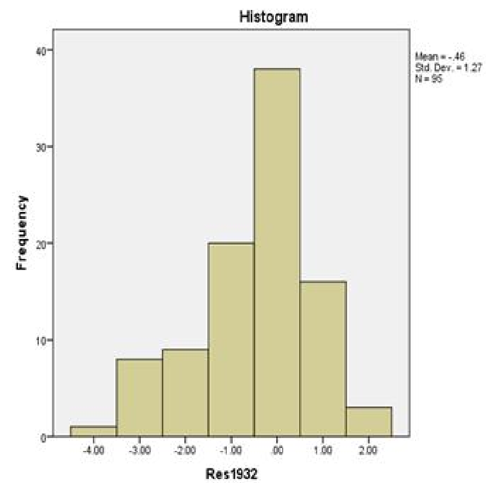
Figure IV: Ordered Regression Residuals

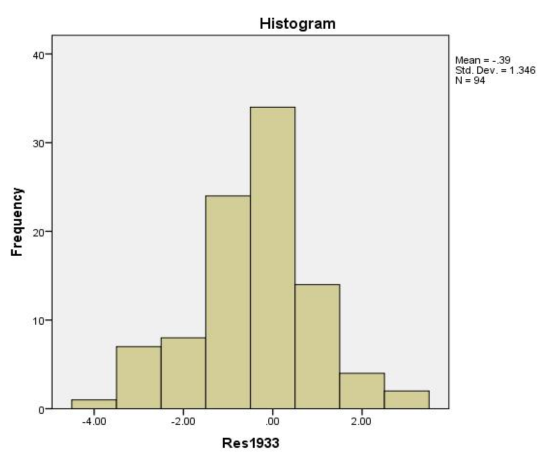












To the extent that volatility remained constant while overall ratings moved down, then the implication is that all ratings declined by approximately the same amount. That is, the residual distribution tended to shift towards lower ratings. As well, the lack of skewness of the residuals over this period suggests that Moody anticipated that higher-rated and lower-rated railroad firms would be impacted similarly by the credit shocks of the Great Depression. Finally, the normality of the fourth moment may indicate that Moody assigned low probabilities to greater-than-2-sigma credit events, including Black Swan events.

The significance of selected cross-sectional results is tested next using a panel regression approach. In particular, the judgment effects resulting in lower overall ratings during the 1931-33 periods are tested. The implication of this result is that Moody accelerated the downgrade of securities during the depression beyond what his statistical ratings model would have suggested.

This section presents the time-series and cross-sectional (panel) approach to test how Moody's rating process changed through the 1925-33 business cycle and, in particular, in response to developments of the Great Depression. The cross-sectional results above suggested that Moody's ratings did change over time, in terms of the relative importance of the two risk factors: security and salability, and in terms of an overall shift towards lower ratings in the 1930s. The panel regression will provide formal tests of these cyclical changes in ratings.

The panel results are summarized in terms of five statistical models. Model 1 presents the basic regression model with variables: (1) salability, which takes values Fair (incorporated with the intercept), Good (SalG), High (SalH), and Very High (SalVH), and (2) TIE which is a continuous measure. The results show that Salability and TIE are highly significant over the entire study period: 1925-33 and that higher levels of both variables lead to higher bond ratings. These two variables constitute Moody's statistical rating.

Model 2 incorporates yearly dummy variables D25-D33 alone without the two regression variables, to test how ratings changed year-to-year without controlling for regression effects. Significant ratings shifts did occur over the study years. On average, ratings improved over 1925-28 (with dummy variables becoming more negative), and overall ratings declined over 1928-33. These rating shifts roughly correspond to the business cycle of the period.

Model 3 incorporates both the yearly dummy variables and the Salability and TIE factors. Salability and TIE remain significant in the present of the yearly dummies. As well, the annual dummy variables continue to show a shift to lower ratings in the 1930s.

In model 4 the yearly dummies are dropped, while interaction affects are added between year and the Salability and TIE variables. In particular, the interaction variables: SalG31_3, SalH31_3, SalVH31_3 and TIE31_3 estimate how the coefficients for these variables might shift during the depression years: 1931-1933. These shifts prove to be significant. The interaction effects for salability are positive and significant, indicating that salability carried less weight in the ratings during the 1930s. In contrast, the interaction effects for TIE are negative and significant, indicating that the weight of this variable increased in the ratings during the 1930s.

Finally, model 5 incorporates both the basic variables Salability and TIE, and the yearly dummies, and the shift dummies. The major conclusions remain intact. Ratings were lower during the 1930s, even after controlling for the decline in salability and TIE variables. The impact of salability, Moody's measure of a security's liquidity, was significant during the 1920s, but declined in importance over 1930-33. In contrast, the solvency measure TIE was significant and increased in importance over 1930-33. This shift in importance perhaps reflects the increasing solvency crisis of the Great Depression.

Reviewing the various models suggests the following.

- The dummy variables: D25 through D33 in Model 3 and 5 suggest a strong overall downgrade trend in railroad bond ratings during the Great Depression
- Comparing models 1, 3, 4 and 5 indicates that the impact of the Salability variable on ratings diminished over these years. That is, a bond's Salability designation of *Good*, *High* or *Very High* would achieve a higher bond rating before 1931, but a lower bond rating between 1931

and 1933. Therefore, during 1931-33 bonds had to achieve a higher level of liquidity to achieve the same bond rating.

- In contrast, the TIE variables in models 1, 3, 4 and 5 indicate a positive association between coverage ratios and bond ratings, and that over 1931_33 the same TIE ratio earned the bond a higher rating by Moody's. That is, the variable TIE31_3 (with coefficient estimates model 4 = -0.169 and model 5 = -0.228) indicates that the variable had a significantly greater impact on boosting ratings during the Great Depression.

Table IV: Data Panel Analysis

Dependent Variable: Moody's Railroad Bond Ratings

	Model 1	Model 2	Model 3	Model 4	Model 5
Threshold parameters:					
"Aaa"	-2.261***	-0.950***	-2.241***	-2.554***	-2.426***
"Aa"	-1.609***	-0.371***	-1.559***	-1.876***	-1.741***
"A"	-1.108***	0.072	-1.023***	-1.337***	-1.192***
"Baa"	-0.571***	0.526***	-0.430**	-0.737***	-0.576***
"Ba"	0.005	1.011***	0.210	-0.080	0.106
"B"	0.674***	1.545***	0.990***	0.693***	0.935***
C-rating	1.119***	1.896***	1.458***	1.115***	1.370***
Location parameters:					
D26		-0.143	-0.233		-0.270
D27		-0.058	-0.162		-0.136
D28		-0.235	-0.345*		-0.340*
D29		0.027	0.040		0.070
D30		-0.074	-0.049		-0.066
D31		0.227	0.292*		0.338
D32		0.552***	0.678***		0.718**
D33		0.732***	0.918***		0.960***
SalG	-0.223*		-0.295**	-0.713***	-0.474***
SalH	-0.823***		-0.939***	-1.343***	-1.120***
SalVH	-0.932***		-1.160***	-1.655***	-1.430***
SalG31_3				1.185***	0.608*
SalH31_3				1.288***	0.743***
SalVH31_3				1.466***	0.829***
TIE	-0.185***		-0.158***	-0.166***	-0.151***
TIE31_3				-0.188***	-0.269***
R^2	87.9%	86.0%	96.2%	94.5%	96.2%

Significance: *=(p=0.05), **=(p=0.01), ***=(p=0.001)

V. Conclusions

This study analyzed Moody's bond rating data using an ordinal regression approach. The rating stability hypothesis states that the rating model and its parameterization should remain constant through the business cycle. In contrast, this study finds that the importance of the security and salability factors shifted, with the solvency variable becoming more important and the liquidity variable becoming less important as the Great Depression became more of a solvency crisis. Thus, ROD would imply that the impact of Moody's judgment would not reflect the nuances of the business cycle. However, the study results reject this hypothesis. Moody's expert judgment lagged the market while he tended to upgrade bond ratings relative to their statistical ratings from 1925 through 1930 and downgrade bonds relative to their statistical ratings from 1931 through 1933.

The Great Depression had two major impacts on Moody's rating process. First, the weight given in the statistical rating to the salability factor declined, while the weight given to the security factor increased. This result is consistent with the Great Depression turning from a liquidity crisis to a solvency crisis as the depression deepened.

Our analysis provides a strong test of the ROD or rating TTC hypothesis. Under this hypothesis Moody's judgment would change only slowly over time, more in response to secular rather than cyclical effects. The data shows, however, that Moody's judgment tended to up-grade bonds on average during the expansion years of the 1920s and downgrade bonds during the depression years of the 1930s. Therefore, his judgment appears to have been strongly procyclical.

Contributions

This paper makes a significant contribution to new knowledge in that analysis was not previously conducted on the expert judgment of John Moody and the Great Depression. Our analysis reflects that Moody downgraded the overall ratings based upon his judgment, above and beyond the downgrading that occurred from the statistical ratings model. If we take this forward to the 2008 market meltdown one can hypothesis if this form of Moody's expert judgment were used we may have avoided the market crash since the rating agencies tended to "support bad debt" (Baily 2009) during the crisis whereas the John Moody analysis tended to be tougher on bad debt. Thus, this paper has presented argument that supports current thinking around past agency ratings as well as presenting new analysis of the 1925-1933 railroad bond ratings of John Moody.

REFERENCES

Altman, E., & Rijken, H. (2005). The Effects of Rating Through the Cycle on Rating Stability, Rating Timeliness and Default Prediction Performance. NYU Working Paper New York, BYU: 10-12.

Amato, J., & Furfine, C. (2003). Are credit ratings procyclical? BIS Working Papers. Monetary/Economic. Washington, D.C., Federal Reserve Bank of Chicago. 129: 1-41.

Baily, M. N., & Elliott, D. (2009). The US Financial and Economic Crisis: Where Does It Stand and Where Do We Go From Here? The Initiative on Business and Public Policy. Washington, D.C., The Brookings Institution: 1 - 26.

Cantor, R., & Packer, F. (1994). "The Credit Rating Industry." FRBNY Quarterly Review Summer-Fall: 1-26.

Kiff, J., Kisser, M., & Schumacher, L. (2013). Rating Through-the-Cycle: What does the Concept Imply for Rating Stability and Accuracy? IMF Working Paper: Monetary and Capital Markets: 20-22.

Loeffler, G. (2004). "An anatomy of rating through the cycle. ." Journal of Banking and Finance 28: 695-720.

Loeffler, G. (2006). Can rating agencies look through the cycle? . Working Paper.: 1-10.

Moody, J. (1909). Key to Moody's Ratings. Moody's analysis of investments - Steam Railroads. New York, Fredrick C Mathieson. 1: 194.

Moody, J. (1913). Key to Moody's Ratings. Moody's analysis of investments - Steam Railroads. F. C. Mathieson. New York. 1: 198.

Moody, J. (1920). Key to Moody's Ratings. Moody's analyses of investments - Steam Railroads. New York, Press of Publishers Printing Company. 12: 27-28.

Shanteau, J., & Stewart, T.R. (1992). "Why study expect decision making? Some historical perspectves and comments." Organizational Behavior and Decision Processes 53: 95-106.

White, L. J. (2010). "Markets: The Credit Rating Agencies." The Journal of Economic Perspectives 24(2): 211-226.

Wilson, B. (2011). "On the information content of ratings: an analysis of the origin of Moody's stock and bond ratings." Financial History Review 18(2): 155-190.

Wolfson, J., & Crawford, C. (2010). "Lessons from the current financial crisis: Should credit rating agencies be re-structured?" Journal of Business & Economics Research 8(7): 85-91.