

MODELING METHODOLOGY

Authors

Jimmy Huang
Mariano Lanfranconi
Nihil Patel
Libor Pospisil

Contact Us

Americas
+1-212-553-1653
clientservices@moodys.com

Europe
+44.20.7772.5454
clientservices.emea@moodys.com

Asia (Excluding Japan)
+85 2 2916 1121
clientservices.asia@moodys.com

Japan
+81 3 5408 4100
clientservices.japan@moodys.com

Modeling Credit Correlations: An Overview of the Moody's Analytics GCorr Model

Abstract

Describing the dependency structure of instruments in a credit portfolio is imperative for measuring credit risk at the portfolio level. The Moody's Analytics credit portfolio framework does not model this dependency structure by making direct assumptions about co-movements of exposure values, but rather focuses on correlations—called asset correlations—among credit quality changes of the corresponding obligors. The Moody's Analytics Global Correlation Model (GCorr™) is a multi-factor model developed by Moody's Analytics which provides such asset correlations. GCorr produces asset correlations of various asset classes at a granular level, and can integrate all asset classes within one model.

This document provides an overview of the GCorr framework, methodology, data used for estimation, and validation. In addition, this document describes the components of GCorr related to individual asset classes and their integration. The asset classes explicitly included in GCorr are: public firms, private firms, small and medium-sized enterprises, sovereigns, U.S. commercial real estate, and U.S. retail. We also discuss how the GCorr model is used to model correlated defaults and recoveries when performing credit portfolio analysis.

Table of Contents

1	Introduction	5
1.1	Importance of Correlations in Modeling Portfolio Loss.....	5
1.2	What is GCorr?	5
1.3	Modeling Credit Correlations.....	6
2	GCorr Framework.....	7
3	GCorr Asset Classes.....	8
3.1	Public Firms.....	8
3.2	Small and Medium Enterprises.....	9
3.3	Sovereigns.....	11
3.4	Commercial Real Estate.....	12
3.5	Retail	13
3.6	Integrating the Models.....	14
4	Correlated Defaults and Recoveries.....	15
5	Validation.....	15
6	Future Extensions	17
	References.....	19

1 Introduction

This section provides information about the importance of correlations, describes the Moody's Analytics Global Correlation Model (GCorr™), and discusses the challenges of modeling asset correlations.

1.1 Importance of Correlations in Modeling Portfolio Loss

Credit portfolio management has long been an area of focus for financial institutions. While all credit portfolios are subject to risk of losses, the uncertainty of losses can be minimized through proper diversification. However, it can be very challenging to accurately model the sources of diversification and pockets of concentrations within a portfolio. Credit risk practitioners often tackle this problem by estimating a granular correlation model that can characterize the level of dependency between the borrowers within their portfolio. Past events have demonstrated that it is imperative to be able to describe the interconnectedness of borrowers within a portfolio. For example, the recent financial crisis has shown the importance of accurately assessing the risks of one asset class (residential mortgages) and then understanding the links with the rest of the economy. Similarly, the varying degrees with which institutions and countries were exposed to the European sovereign debt crisis highlighted the need for correlation models that can precisely characterize the level of dependency. The inability to properly model the various risks of a portfolio has led to billions of dollars in losses in some cases, and has resulted in the failure of banks that did not hold enough capital to sustain those losses.

Moody's Analytics Economic Capital (EC) platform, RiskFrontier®, employs a bottom-up approach to estimate a portfolio value distribution at a future time horizon.¹ The portfolio analysis starts with modeling each individual borrower by running a Monte Carlo simulation that generates systematic shocks and borrower-specific shocks. The systematic shocks describe the future state of the economy, and the borrower-specific shocks model the independent risks faced by each borrower. Together, these simulated shocks establish a borrower's credit quality in the future. Since all borrowers are exposed to a set of correlated factors, the credit quality changes across borrowers are correlated. A valuation framework is then applied to model the value of every instrument based on the future credit quality of the borrower, and the future value distribution of the portfolio can then be constructed using a large number of trials.

1.2 What is GCorr?

GCorr is a model produced by Moody's Analytics which is used to estimate correlations among credit quality changes of obligors in a credit portfolio. GCorr includes correlation estimates for a variety of asset classes that are extensively validated to ensure the correlations produced by the model are in line with real world patterns and economic experience. The factor framework in GCorr is also versatile in that it provides correlations between any two borrowers across a wide range of asset classes. The model is granular enough to differentiate the risks of borrowers within and across a variety of risk factors.

Financial institutions often hold exposures across a variety of asset classes. To properly estimate the value distribution of the portfolio, it is essential to use a correlation model that captures the specific behavior of each asset class given the empirical variation in correlations. The same model for every asset class does not capture the rich cross-sectional variation in correlations. For example, corporate firms in the same region and industry are often exposed to similar risks, so their correlations are relatively high. On the other hand, individual retail borrowers have much lower correlations, since one credit card borrower defaulting will likely have little impact on another credit card borrower's ability to pay. Because asset classes exhibit different behaviors, rigorous models must be calibrated for each one.

¹ RiskFrontier provides a single platform and analytical solution for modeling portfolio credit risk, integrating all the relevant risk factors across a wide variety of asset classes to produce a holistic measure of risk, expressed as Economic Capital. For more information, see "An Overview of Modeling Credit Portfolios" (Levy, A. 2008).

Not only is it important to accurately model correlations within each asset class, one must also consider the inter-correlations across asset classes. Certain asset classes show significant levels of correlations with each other, and this effect cannot be ignored. For example, many institutions have exposure to the sovereign debt of troubled European countries. With the European sovereign debt crisis, one sovereign defaulting could negatively affect the credit quality of many institutions. Therefore, it is important to be able to capture such correlations. GCorr provides a flexible framework that can model such correlations across asset classes.

Another benefit of GCorr is the ability to differentiate between the risks of borrowers across several factors. For example, a less detailed correlation model might not be able to differentiate the risks of two auto loans originated in neighboring cities, but the retail correlation model in GCorr provides different risk measures for 250 major U.S. cities across six different product types. This risk measure is further differentiated by whether the borrower is prime or sub-prime. Such detailed classification allows for a better understanding of the exposures that are contributing the most towards the overall portfolio risk.

To summarize, GCorr is a granular model that estimates the forward-looking correlations between any two borrowers. Currently, there are models covering the following types of borrowers:

- Public corporate firms
- Private firms
- Small and medium enterprises (SME)
- U.S. retail
- U.S. commercial real estate (CRE)
- Sovereigns

GCorr is designed with flexibility in mind. Its structure allows for an arbitrary extension of customized factors so that it can be fine-tuned to a financial institution's needs. In addition, utilizing the flexible correlation framework of GCorr, Moody's Analytics has created bespoke correlation models calibrated using client specific data across the globe. Finally it is worth mentioning that the bottom-up simulation framework employed in RiskFrontier allows for the credit portfolio analysis to model dynamics associated with a correlation structure between defaults and recoveries. In this document we describe how RiskFrontier models correlated defaults and recoveries and how we parameterize these correlations.

1.3 Modeling Credit Correlations

There are several measures of correlation to consider when modeling portfolio credit risk. Value correlation measures how the values of two instruments co-move. Default correlation describes how the default of one borrower relates to the default of another. If only the default/no-default states are considered at horizon, the value correlation is equal to the default correlation. Because the Monte Carlo framework within RiskFrontier simulates the future credit state of each borrower and models the instrument's value based on the credit state, GCorr focuses on the estimation of credit quality co-movements of borrowers. Note that the credit quality co-movements of borrowers are known as *asset correlations* within the GCorr framework.

Each asset class has its own behavior and available data, so there is no generic estimation methodology that can be applied across all asset classes. For corporate firms, we estimate credit quality co-movements by measuring how the asset values of two firms co-move within the Vasicek-Kealhofer Model. For sovereigns and CRE borrowers, we leverage probability of default (PD) models to estimate a proxy for credit quality co-movements. For SME and retail loans, we use default data to calibrate the correlation parameters. For an overview of how correlations are modeled in each of the asset classes, see Section 3.

2 GCorr Framework

GCorr utilizes a multi-factor model to estimate asset correlations of borrowers. The asset return of each borrower is affected by a systematic factor and an idiosyncratic factor:

$$r_i = \sqrt{RSQ_i}\phi_i + \sqrt{1 - RSQ_i}\epsilon_i \quad (1)$$

r_i is the asset return of borrower i

ϕ_i is the systematic factor return of borrower i

RSQ_i is the R-squared of borrower i – the proportion of risk that is captured by the systematic factor

ϵ_i is the idiosyncratic shock of borrower i

The systematic factor represents the state of the economy and summarizes all the relevant systematic risk factors that affect the borrower's credit quality. Within GCorr, the systematic factor ϕ_i is a weighted combination of 245 correlated geographical and sector risk factors where the weights can be unique to each borrower. The idiosyncratic shocks represent the borrower-specific risk that affects the borrower's credit quality. While borrowers with the same weights to the 245 factors are exposed to the same systematic shock, the borrower-specific shock is unique to each borrower. By construction, the systematic factor is independent of the idiosyncratic factor and both are modeled with a standard normal distribution. Two borrowers correlate with one another when both are exposed to correlated systematic factors.

The correlation between changes in credit quality for any two borrowers, both within and across asset classes, is equal to:

$$\begin{aligned} corr(r_a, r_b) &= corr(\sqrt{RSQ_a}\phi_a + \sqrt{1 - RSQ_a}\epsilon_a, \sqrt{RSQ_b}\phi_b + \sqrt{1 - RSQ_b}\epsilon_b) \\ &= \frac{cov(\sqrt{RSQ_a}\phi_a + \sqrt{1 - RSQ_a}\epsilon_a, \sqrt{RSQ_b}\phi_b + \sqrt{1 - RSQ_b}\epsilon_b)}{\sigma_{r_a}\sigma_{r_b}} \\ &= \frac{\sqrt{RSQ_a}\sqrt{RSQ_b}corr(\phi_a, \phi_b)}{\sqrt{1 \times 1}} \\ &= \sqrt{RSQ_a}\sqrt{RSQ_b}corr(\phi_a, \phi_b) \end{aligned} \quad (2)$$

Therefore, the only input parameters required to calculate the correlation between any two borrowers are the R-squared values of the borrowers and the correlation between their systematic factors.

This factor framework has several advantages over calculating pairwise empirical asset correlations directly from data:

- The amount of information that needs to be retained is dramatically reduced. For a portfolio of 1,000 assets, almost half a million pairwise correlation estimates must be stored. With the GCorr factor approach, this reduces to 1,000 R-squared estimates and the covariance matrix of 245 factors.
- Pairwise correlation estimates are extremely noisy, and the confidence interval around correlations can be very large. The factor model approach isolates the systematic component from the noise, thus modeling the correlation based solely on the exposure to the systematic factors and mitigating the impact of the noise.
- The data for some borrowers may not be sufficient. With the factor framework, as long as an R-squared value can be modeled for the borrower, the correlation with any other borrower can be computed. Without a factor model, one would require a sufficiently long time series of asset returns.

3 GCorr Asset Classes

This section discusses each individual GCorr asset class. It describes the available data, the general methodology used to estimate the R-squared values and systematic factors, and certain modeling decisions.

3.1 Public Firms

GCorr Corporate is the global correlation model for publicly traded firms. Banks typically use the model to determine the correlations among borrowers in their Commercial and Industrial (C&I) portfolios.

In GCorr Corporate, the systematic risks of firms are assumed to depend on countries and industries within which firms operate. Therefore, the systematic factors in this model are computed as a weighted sum of country factors and industry factors. When we estimate the model, the country factor of each firm is determined by the country of incorporation, and the industry component is a weighted combination of industry factors in which it operates, such that the industry weights sum to 1. There are a total of 110 country and industry factors (49 country factors and 61 industry factors) estimated in GCorr Corporate, which covers 2,989 (49 x 61) country-industry categories. The two outputs from the model are the firm-level R-squared values and the covariance matrix of the 110 country and industry factors.

An important modeling decision is the length of time used to estimate R-squared values and systematic factor correlations. One interesting observation is that R-squared values of firms are cyclical. During bad economic times, firm credit qualities tend to move closer with the systematic factors and therefore are more sensitive to the state of the economy, resulting in higher R-squared values. We also find that periods of high systematic factor variance are associated with periods of high asset correlations.² The time period used to estimate a firm's R-squared should be short enough to capture these dynamics, but at the same time should be estimated with enough observations to have a robust estimate. We find that using a 1-year window for R-squared estimation leads to R-squared values that are volatile over time. Instead, we use a 3-year window, which provides smoother R-squared values from year to year, but still captures the cyclical dynamics of the business cycle. To calculate the systematic factor correlations, we want to use a long enough period to capture at least one full business cycle. In addition, we find that the variation in systematic factor correlations is not as pronounced as the variation in R-squared values.

GCorr Corporate is a forward-looking correlation model. Due to the cyclical nature of R-squared values and correlations, GCorr combines short-term and long-term components to predict how correlation levels will move in the future. This is important because correlations change through time, and to have a proper model for the future portfolio value distribution, one must consider how correlations will move in the future.

3.1.1 Data and Methodology

The GCorr Corporate model uses historical weekly asset returns of firms around the world starting from July 1, 1999. Asset returns are computed based on the Vasicek-Kealhofer framework that considers equity as a call option on assets.³ The asset value is therefore mainly driven by equity prices along with debt information (amount of debt, duration, interest rate level/change). It is worth noting that for highly levered firms—which also tend to be the riskiest—we find notable differences between equity and asset return correlations. That is, equity correlations are not a good proxy for asset correlations in some cases.

Before doing any computations, we thoroughly clean the data to remove both outliers in a firm's time series of asset returns and outliers across all the firms at each time step. The credit risk of a large corporate firm is mainly driven by the region and industry within which the firm operates, so the systematic factor is defined as a combination of country and industry factors. To estimate the country and industry factors, market-weighted asset return indexes are computed for

² For more information, see "Understanding Asset Correlation Dynamics for Stress Testing" (Cai, Q., A. Levy, and N. Patel 2009).

³ The credit quality co-movements of borrowers are known as *asset correlations* within the GCorr framework

each country and industry combination with a sufficient number of firms. The return for a country-industry combination is further decomposed into a country return and an industry return through a regression. This methodology essentially uses data from countries and industries with rich data to infer factor returns for countries and industries with limited data. For firms without a sufficient amount of equity returns, such as newly issued firms, we estimate an empirical model to determine a firm's R-squared value based on its size, country, and industry weights. This Modeled R-squared value is also used to estimate the R-squared values for large private firms. For more information, see Section 3.1.2.

After we estimate the 49 country and 61 industry factors, we can determine the systematic factor for every firm. We then estimate the R-squared values by regressing a firm's asset returns on its systematic factor returns, and make further adjustments to ensure that the R-squared values forward-looking. Because the time series of all the systematic factors can be estimated, it is also straightforward to estimate the correlations between the systematic factors. Having computed an R-squared value for each firm and the correlation between all the systematic factors, we can then easily compute the correlation between any two corporate borrowers.

3.1.2 Modeled R-Squared

For corporate firms with little or no equity data, such as firms that recently had an initial public offering or large private firms, we can leverage information from publicly traded firms that have rich equity data to estimate Modeled R-squared values. The Modeled R-squared estimation begins with the observation that there is a strong positive relationship between the R-squared value of a firm and its size. Large firms tend to be more sensitive to the state of the economy and often have a larger systematic exposure, so we see higher R-squared values. On the other hand, small firms are less influenced by the economy, resulting in smaller R-squared values. After controlling for firm size, we find cross-sectional differences in R-squared across countries and industries.

To estimate this model, R-squared values are first estimated for firms with a sufficient amount of equity data. Due to the strong relationship between R-squared values and size, we run a two step regression. In the first step, we regress R-squared values on size. In the second step, we regress the residuals from the first step on country and industry dummy variables to determine how much of the residual is explained by the country and industry components. Using the results of these regressions, we can estimate the Modeled R-squared value based on the size, country, and industry weights.

The model's advantage is the granularity of R-squared values. It can estimate Modeled R-squared values for any country, industry, and size, allowing for differentiation across the various exposures in a portfolio. Institutions with exposures to private firms will typically use Modeled R-squared values. However, because Modeled R-squared values are calibrated with publicly traded firms, this model should only be used for large private firms which typically have access to financing via the capital markets. For small and medium enterprises, Moody's Analytics has developed a specific correlation model. Private corporate firms are found to be empirically exposed to similar systematic components as large publicly traded firms, so the systematic factors are the same as in the Gcorr Corporate model. With the Modeled R-squared values and the correlation between systematic factors, one can compute the correlations between private borrowers, as well as the correlation between a private borrower and a borrower in another asset class.

3.2 Small and Medium Enterprises

In general, Small and Medium Enterprises (SMEs) exhibit different correlation patterns than larger firms.⁴ In particular, their correlation levels are lower than those of large corporates. Since the Modeled R-squared is estimated using publicly traded firms that are much larger than SMEs, that model might not provide reasonably accurate correlation estimates for SMEs. Fortunately Moody's Analytics is able to leverage a propriety database of private firm financial statements and

⁴ Small and Medium Enterprises (SMEs) can be defined in various ways, either by the number of employees or the amount of revenue. In Europe, SMEs are defined as firms with revenue of €10-50 million. Individual countries can have their own definitions as well. For example, Germany classifies a company as an SME if it has fewer than 255 employees, whereas in Belgium the limit is 100.

defaults to calibrate correlations for SMEs. Currently, there are models estimated for five countries: United States, United Kingdom, France, Canada, and South Africa.

Due to the lack of information available for individual SMEs, the estimation methodology groups SMEs into homogeneous pools for each industry and region combination. Pooling the firms allows us to compute a default rate time series based on available data. From the dynamics of the default rate, we can estimate the implied R-squared value for each pool using the method of moments. The SME model also uses the same systematic factors as the corporate model, defined as a weighted combination of country and industry factors. The R-squared values and the systematic factor correlations allow for the estimation of the correlation between any two SME firms.

To determine if a private firm should be modeled using the Modeled R-squared or the SME model, Moody's Analytics provides a size threshold to indicate which model is appropriate. If a firm is larger than the size threshold, the Modeled R-squared should be used because it was estimated using data from large corporate firms. If a firm is underneath the size threshold, it is recommended to use the SME model.

3.2.1 Data and Methodology

Moody's Analytics uses its own Credit Research Database (CRD) to estimate R-squared values for SMEs. The CRD is built in partnership with leading financial institutions around the world to create a data source of private firm financial statements and private firm defaults. For the United States, the data is granular enough to pool into ten regions and 61 industries (610 pools). For the other countries, data is pooled into 61 industries.

To estimate the R-squared values, firms in the CRD are pooled by country/region and industry. The dynamics of the realized default rate are used to estimate the implied R-squared value for each pool using the method of moments.

Because all firms in the homogeneous pool share the same systematic factor and R-squared, the asset correlation between any two firms in the pool will be equal to the R-squared of each borrower in the pool. The estimation methodology starts with the distribution of the aggregate default rate for a group of loans based on the underlying PD and asset correlation RSQ of the firm pairs in the group.

Consider a homogeneous group of N unlisted firm loans with the identical default probability of PD and asset correlation parameter RSQ . Let L_i be the indicator for the realization of default for account i , and let L be the percentage of defaults (i.e., default rate) in the pool during a particular time period:

$$L = \frac{1}{N} \sum_{i=1}^N L_i \quad (3)$$

Following the setup in Vasicek (2002), when the number of loans is infinitely large and granular, the probability distribution converges to a limiting form:

$$\text{Prob}(L \leq x) = N\left(\frac{\sqrt{1 - RSQ}N^{-1}(x) - N^{-1}(PD)}{\sqrt{RSQ}}\right) \quad (4)$$

The mean and variance of the default rate for this group are as follows:

$$\begin{aligned} E(L) &= PD \\ \text{Var}(L) &= N_2(N^{-1}(PD), N^{-1}(PD), RSQ) - PD^2 \end{aligned} \quad (5)$$

For a given homogeneous group of unlisted firms, we can calculate the sample mean $\hat{\mu}$ and sample variance \hat{s}^2 of the realized default rate time series as the estimates of PD and $Var(L)$, respectively. Then, we can back-out an R-squared estimate \widehat{RSQ} for the pool by solving the following equation:

$$\hat{s}^2 + \hat{\mu}^2 = N_2(N^{-1}(\hat{\mu}), N^{-1}(\hat{\mu}), \widehat{RSQ}) \quad (6)$$

The solution to this equation, \widehat{RSQ} , is called the method of moments estimator (MME) of R-squared, which refers to the fact that Equation (6) matches the empirical and theoretical moments of the default rate distribution.⁵ From the method of moments estimator, we adjust the R-squared values to control for different factors that may bias these values. We find that the results are biased by several factors, such as the number of borrowers in a pool, the number of time periods, the degree of heterogeneity, and the autocorrelation in the default rate series. We then determine the effect of the bias for each of these variables and remove these effects to arrive at the refined R-squared estimates.⁶

3.3 Sovereigns

To model a sovereign entity's credit quality changes, we leverage the CDS markets. The CDS spread allows us to back out the CDS-implied Distance-to-Default (DD).⁷ Changes in CDS-implied DD can be considered proxies to credit quality co-movements, so the time series of CDS-implied DD changes are used to determine the R-squared values.

The systematic factor of a sovereign is defined as a weighted combination of country and industry factors estimated from the GCorr Corporate model. This allows us to estimate the correlation between sovereigns and other asset classes. The country component is computed as a weighted average of the country factors within the sovereign's region, with additional weight placed on the sovereign's own country. The industry factor weights reflect the contributions of industries to the GDP of the sovereign's region. In the corporate model, each firm loads only to its country of incorporation, but sovereigns can also be affected by the countries around them. For example, the credit risk of many European countries is dependent on the health of the European economy. Since many countries hold debt from other sovereigns, one sovereign defaulting can greatly impact the credit risk of other sovereigns. Additionally, we see that loading to countries within the sovereign's region leads to a greater consistency of the model, resulting in a better fit with empirical patterns in the data.

3.3.1 Data and Methodology

The sovereign correlation model uses CDS spreads provided by Markit. Before any estimation, we filter out any illiquid CDS spreads. From the observed CDS spreads, we can compute the CDS-implied DD changes for each sovereign. Because CDS-implied DD changes are a proxy for credit quality co-movements, we estimate the empirical correlations between the sovereigns by calculating the correlation of CDS-implied DD changes. To apply the GCorr framework, we must estimate the R-squared value of each sovereign.

Our empirical work finds that the systematic factors of each sovereign can be computed as a weighted combination of the country and industry factors from GCorr Corporate. This allows us to calculate the correlation between every pair of

⁵ In addition to the method of moments presented in Section 3.2.1, there are other methods for estimating R-squared values from default rate time series. For more information, see "A Comparative Anatomy of Credit Risk Models" (Gordy 2000) and "Estimating Default Correlations from Short Panels of Credit Rating Performance Data" (Gordy and Heitfield 2002).

⁶ For more information, see "Modeling Correlations of Small and Medium Enterprises" (Huang J., N. Patel, and L. Pospisil 2012).

⁷ Distance-to-Default (DD) is a measure produced by the Moody's Analytics Public Firm EDF Model. The DD is the difference between the expected market value of assets and the default point, divided by the volatility of the asset returns.

sovereign's systematic factors. We can then use the empirical CDS-implied DD correlations from the previous step to imply the R-squared values.⁸

$$\begin{aligned} \text{corr}(r_a, r_b) &= \sqrt{\text{RSQ}_a} \sqrt{\text{RSQ}_b} \text{corr}(\phi_a, \phi_b) \\ \sqrt{\text{RSQ}_a} \sqrt{\text{RSQ}_b} &= \frac{\text{corr}(r_a, r_b)}{\text{corr}(\phi_a, \phi_b)} \end{aligned} \quad (7)$$

where $\text{corr}(r_a, r_b)$ is the empirical CDS-implied DD correlation and $\text{corr}(\phi_a, \phi_b)$ is the systematic factor correlation of the sovereigns. The equation above provides a method for estimating the R-squared values of sovereigns with CDS data of sufficient quality. In addition, we build an empirical model that links these estimated R-squared values to sovereign characteristics, such as region, credit quality, and GDP. The model makes it possible to imply R-squared values for sovereigns without CDS data of sufficient quality.

3.4 Commercial Real Estate

Commercial Real Estate (CRE) exposures often represent a large share of the credit portfolios held by many financial institutions. Therefore, it is essential to accurately model the correlations associated with these loans, as well as the correlations with the rest of the portfolio. GCorr provides a CRE framework that models the correlations of different types of CRE loans in the United States.

Modeling the correlations between CRE loans can be challenging. We can compute asset returns for large corporate firms, but this is not possible with CRE loans. Ideally, we would prefer to use a commercial property's total return time series because they are similar in nature to asset returns. However, property level capitalization rates are not currently available, so we rely on parameters estimated from dynamics of property level net operating income (NOI) in order to capture idiosyncratic component of a CRE borrower's credit risk.⁹

Because CRE loan credit risk is affected by the city and property type, the systematic factors in this model are computed as a sum of a Metropolitan Statistical Area (MSA) factor and a property type factor. The model has 73 MSA factors and the following five property type factors:

- Hotels
- Industrial
- Multi-Family Housing
- Office
- Retail

In the CRE framework, we group CRE loans into MSA and property type pools and consider them to be homogeneous. In other words, the CRE loans in the same MSA and property type combination have equal exposures to the same systematic factors, so they have the same R-squared value. Our data is granular enough to produce 365 R-squared values for every MSA and property type combination.

The model's level of granularity is already a substantial improvement over alternative models, which tend to be based on a coarser classification. This parameterization enables substantially more accurate portfolio credit risk analysis. Currently, one drawback of the GCorr CRE model is that it is only applicable for U.S. CRE portfolios. As we collect more sufficient and robust CRE data internationally, similar correlation models can be built for other parts of the world.

⁸ For more information, see "Modeling Sovereign Correlations" (Levy, A., N. Patel, L. Pospisil, and V. Sesum 2011).

⁹ A CRE borrower refers to the underlying collateral backing a CRE loan (i.e., the commercial property).

3.4.1 Data and Methodology

The data used for estimating the GCorr CRE model comes from Moody's Analytics CMM® (Commercial Mortgage Metrics) model.

In the CMM framework, the underlying collateral value backing a CRE loan (the commercial property) is modeled as a stochastic process driven by both market-wide and idiosyncratic factors. Monte Carlo simulation is used to generate future paths of collateral NOI and market value. In the model, a CRE loan credit event is doubly triggered by the collateral's financial condition at the time of default: both the sustainable NOI falls below the total debt service, and the property's market value falls below the total outstanding loan balance. To capture the actual observed borrower default behavior, CMM is empirically calibrated to historical defaults. CMM uses the following datasets to estimate collateral NOI and value:

- National Council of Real Estate Investment Fiduciaries (NCREIF)
- CBRE Econometric Advisors
- Real Capital Analytics (RCA)
- Trepp's CMBS Deal Library

For each MSA and property type combination, we estimate a commercial property's R-squared value by measuring the proportion of the property's NOI variation explained by the market level NOI variation.

To estimate the systematic factors, we use time series of default probabilities from CMM. Each time series represents default probabilities for a typical property in a given market. The time series variation in the default probability is driven by market-wide variables and therefore does not depend on any idiosyncratic components. Based on the probability of default (PD), we can determine the PD-implied Distance-to-Default (DD) of the typical CRE borrower in a given market. As mentioned in the sovereign section, the PD-implied DD changes are a proxy for credit quality movements. We use the changes in PD-implied DD for all the MSA and property type combinations that are available to construct the 73 MSA and five property type factors. Finally, we can combine every combination of MSA and property type factors to form the 365 systematic factors.

3.5 Retail

Similar to the U.S. CRE model, the biggest challenge in modeling U.S. retail asset correlations is determining an appropriate proxy for systematic factors. It is difficult to model the asset returns of individual borrowers with a high degree of accuracy. Instead we focus on delinquency rates of homogeneous retail pools, where the borrowers in each pool have the same credit risk properties. Similar to the methodology described for SMEs, we can look at the dynamics of delinquency rates and compared them to the expected default rate to model the R-squared values and time series of systematic factors. Having estimated these values, we can then compute all the correlations between retail borrowers and cross correlations between retail and other asset classes.

There are other available measures that can represent the performance of a retail pool as well, such as the charge off rate and the foreclosure rate. The charge off rate is the proportion of accounts charged-off by creditors based on accounting standards, but this measure sometimes suffers from accounting distortions. The foreclosure rate may also misrepresent the performance of a particular retail pool because foreclosure processes usually last several periods and may vary from state to state given the different legislation. As a result, we use the delinquency rate, as it represents a truer performance of the retail pool over any given period.

The retail model produces R-squared values at a very granular level. The model produces an R-squared value for any combination of 300 regions (250 major cities and 50 U.S. states), the borrower credit quality (prime or sub-prime as defined by the credit score), and the following six product types:

- Auto Loan
- Bank Card
- Consumer Loan
- First Mortgage
- Home Equity
- Student Loan

We define the systematic factor as the sum of a state factor and product type factor and use the delinquency rate series to estimate 51 state factors (50 U.S. states and Washington DC) and the six product types.

3.5.1 Data and Methodology

The model is estimated using quarterly delinquency rate time series obtained from Credit Forecast 4.0, a product of Equifax® and Moody's Economy.com. Credit Forecast 4.0 is a population loan level data sourced by vintage, product, origination risk score, and current risk score. The Credit Forecast 4.0 dataset begins in the third quarter of 2005.

Similar to the SME model, we use the method of moments to estimate the R-squared values. As documented in Section 3.2, we essentially solve for the level of R-squared values that matches the first two moments of the delinquency rate time series.

To compute the systematic factors, we assume that the realized default rate DR_t in each pool is equal to the conditional default probability of the underlying borrowers given the systematic shock ϕ_t . Therefore, if we know the R-squared value, we can back out the return on the systematic factor:

$$\phi_t = \frac{1}{\sqrt{RSQ}} \left(N^{-1}(PD_t) - \sqrt{1 - RSQ} N^{-1}(DR_t) \right) \quad (8)$$

After computing the systematic factor returns for all the state and product type pools, we run a regression to decompose the systematic factor return into a state factor return plus a product type return. The result is 51 state factors (50 U.S. states and Washington DC) and six product type factors, which can then be recombined to compute all the 51 x 6 systematic factor returns. While it was possible to compute the systematic factors using MSA level delinquency series, we found that it was sufficient to compute state level factors since MSAs within a state had high levels of correlation.

3.6 Integrating the Models

The benefit of GCorr is its ability to estimate correlations not only within asset classes, but across asset classes as well. This allows financial institutions to compute the correlations between any two exposures in their portfolio, regardless of the asset class. As a result, GCorr can be used to assess the credit risk of a portfolio which includes exposures from various asset classes. This section describes how the GCorr components are integrated.

In the previous sections, we introduced models for the individual asset classes. Each model provides the R-squared values and the systematic factors for each borrower. The corporate model has 110 factors (49 country and 61 industry), the U.S retail model has 57 factors (51 state and 6 product types), and the U.S CRE model has 78 factors (73 MSAs and 5 property types) for a total of 245 correlated systematic factors. The private firm, SME, and sovereign models all share the same systematic factors as the corporate model.

Using the 245×245 covariance matrix of the systematic factors, one can compute the correlation between the systematic factors of any two borrowers. Combined with the R-squared values estimated from each of the individual models, we can now estimate the asset correlation between any two borrowers using the equation:

$$\text{corr}(r_a, r_b) = \sqrt{RSQ_a} \sqrt{RSQ_b} \text{corr}(\phi_a, \phi_b) \quad (9)$$

4 Correlated Defaults and Recoveries

Overwhelming evidence shows that recovery in a default event is closely related to macroeconomic conditions.¹⁰ Generally speaking, recovery is pro-cyclical; during a recessionary period, recovery tends to be lower than during an expansionary period. This positive correlation between recovery and general economic conditions can be rationalized using a demand and supply argument—when a firm is in financial distress and needs to sell assets, its industry peers are likely to experience problems as well. This implies that more distressed assets are available in the market, but fewer firms are able or willing to buy those assets. This supply increase and demand decrease drives the price of the distressed assets below their value in best use.

RiskFrontier accounts for PD-LGD correlation in credit portfolio analysis.¹¹ The model utilizes Moody's Analytics LossCalc™ and GCorr data to overcome challenges in parameterization. Specifically, Moody's Analytics provides the correlation parameters for public and private firms used to calibrate LossCalc. For other asset classes such as CRE and retail, Moody's Analytics works with clients' specific default and recovery data to estimate the PD-LGD correlations parameters.

With this functionality, users can construct an integrated correlation structure that models not only correlation between credit quality co-movements, but also correlation between LGD and credit quality changes of the borrower and the correlation between LGD values of different instruments. The model accounts for the impact on portfolio risk associated with the compounding effects of severe losses during periods when defaults are concentrated. In addition, it accounts for the impact of systematic risk in recovery on valuation and returns. The integrated correlation structure provides portfolio managers the ability to minimize recovery risk in their portfolio through diversification.

At the portfolio level, the values of defaulted instruments are correlated with systematic factors. An implication is that during an economic downturn, not only will the number of defaults be higher, defaulted instruments tend to realize a lower recovery amount as well. As a consequence, portfolio value distribution will have a heavier tail, resulting in higher risk measures (for example, Unexpected Loss and capital) when accounting for PD-LGD correlation. Spreads will also widen to compensate for the increased risk.

5 Validation

A prudent risk practitioner needs to make sure the models are validated. While simplifying assumptions are often used in the model estimation, it is important that the final estimates of the model are reasonable for its intended business use. In the case of credit quality changes, we want our model to provide forward-looking estimates of correlations that can be used within an Economic Capital platform such as RiskFrontier to accurately measure the risk inherent in a credit portfolio.

¹⁰ For a review of literature and empirical evidence, see “Default Recovery Rates and LGD in Credit Risk Modeling and Practice: An Updated Review of the Literature and Empirical Evidence” (Altman 2006).

¹¹ For more information, see “Implications of PD-LGD Correlation in a Portfolio Setting” (Meng, Q., A. Levy, A. Kaplin, Y. Wang, and Z. Hu, 2010) and “Incorporating Systematic Risk in Recovery: Theory and Evidence” (Levy, A. and Z. Hu, 2007).

Every year, various validation analyses are conducted to assess whether the correlation estimates produced by the models are economically intuitive. In the corporate model, we form various groups of firms (500 largest, 1000 random, etc.) for different regions (Europe, North America, etc.) and compare the GCorr modeled correlations with the historical correlations. We test the in-sample correlation by comparing the latest GCorr estimates with the historical data used to calibrate the model. We also test the out-of-sample correlation by comparing the GCorr correlations estimated in previous years with more recent empirical correlations observed in the data. We see that GCorr performs well both in-sample and out-of-sample.

While it is important that the model correlations perform well in-sample and out of sample, we also would like to test model performance using alternative data sources. To this end, Moody's Analytics has conducted validation studies which assess the reasonableness of the correlation estimates when using RiskFrontier.

For example, in one study, Moody's Analytics tested the assumption of using a Gaussian copula to model the joint distribution of asset returns by comparing modeled default correlations implied by GCorr asset correlations with empirical default correlations observed in Moody's Analytics default database.¹² We find that GCorr asset correlations provide a more conservative measure in the sense that they imply higher default correlations than the empirically observed default correlations. Thus, the Gaussian copula together with the GCorr asset correlation levels appears to sufficiently capture the tail risk. It is worth mentioning that we have tried modeling asset correlations using a non-Gaussian copula, but found that the estimated parameters are unstable through time. Ultimately, we find that the Gaussian copula offers mathematical conveniences, provides stable estimates, and is sufficient in capturing the risks of a portfolio.

We conducted another study to determine ex-ante risk measures (unexpected losses, or ULs) of various CDS portfolios and compare them to out-of-sample empirical ULs.¹³ We want to know to what extent the ex-ante ranking of portfolios by ULs predicts the ranking based on future empirical ULs. The results of this exercise depend on the input parameters for the ex-ante ULs, including correlations. The study concludes that using GCorr Corporate correlations results in ex-ante ULs which perform better in identifying more and less risky portfolios than alternative correlation measures, such as empirical asset correlations.

In another study, we compare the realized default distribution to the predicted default distribution using a variety of PD and correlation inputs.¹⁴ The comparison looks at portfolio performance using two sets of estimates: (1) agency ratings-based long-term average default rates with Basel II correlations; and (2) public EDF credit combined with correlations from GCorr. The study finds that the first set of estimates produces a less conservative view of economic capital, whereas the second set produces consistent and conservative economic capital estimates over time.

For the CRE and Retail models, we compare the GCorr correlations with the correlations implied by the default rate time series to validate the GCorr correlations. For the SME model, we look at R-squared values extrapolated from the Modeled R-squared to determine how they compare to the levels of SME R-squared values. For the sovereign model, we test the robustness of the correlation methodology by ensuring that we get the correct regional and overall sample correlation levels.

¹² For more information, see "Asset Correlation, Realized Default Correlation, and Portfolio Credit Risk" (Zhang, J., F. Zhu, and J. Lee 2008).

¹³ For more information, see "Navigating Through Crisis: Validating RiskFrontier® Using Portfolio Selection" (Hu, Z., A. Levy, and J. Zhang 2010).

¹⁴ For more information, see "Economic Capital Model Validation: A Comparative Study" (Hu, Z., A. Levy, and J. Zhang 2012).

6 Future Extensions

Every year, Moody's Analytics conducts research to continually update and improve GCorr. For example, we collect more data and the number of firms used to estimate the corporate model continues to increase. Currently, we estimate GCorr CRE and GCorr Retail with data from the United States only. In the future, if international data becomes available with a sufficient amount of granularity, the models might be extended to cover international markets as well. Similarly, the SME model is gradually expanding to cover more countries. In addition to expanded data coverage, we continue to improve the methodology and make adjustments to achieve the best performance.

Financial institutions often have their own data that can be used to calibrate correlations. Moody's Analytics has worked with numerous clients across the globe to integrate their data with our existing data series to build custom correlation models which reflect the client's customer base.

Most recently, another area of increasing interest is the effect of macroeconomic variables on credit portfolio losses. For example, risk practitioners are trying to understand how movements in macroeconomic variables, such as a large negative shock to the GDP growth, affect the expected loss on a portfolio. To help answer such questions, Moody's Analytics is expanding the GCorr model to include macroeconomic variables. When employed within the RiskFrontier framework, this expanded GCorr model can serve multiple functions. It provides economic interpretation for the GCorr systematic credit risk factors, helps characterize relationships between portfolio losses and macroeconomic variables, and facilitates stress testing and reverse stress testing analyses. In the context of this document, stress testing refers to a calculation of conditional portfolio losses given a macroeconomic scenario, while reverse stress testing describes macroeconomic scenarios associated with a specified level of losses. In addition, the expanded GCorr model can be used to integrate credit risk and other types of risk, such as market risk.

Acknowledgements

We would like to thank Amnon Levy and Jing Zhang for research guidance and comments.

Copyright © 2012 Moody's Analytics, Inc.
and/or its licensors and affiliates. All rights reserved.

References

- Altman, E., 2006, "Default Recovery Rates and LGD in Credit Risk Modeling and Practice: An Updated Review of the Literature and Empirical Evidence," New York University Salomon Center, Stern School of Business.
<http://pages.stern.nyu.edu/~ealtman/UpdatedReviewofLiterature.pdf>
- Cai, Q., A. Levy, and N. Patel, 2009, "Understanding Asset Correlation Dynamics for Stress Testing." Moody's Analytics White Paper.
- Gordy, M. B., 2000, "A Comparative Anatomy of Credit Risk Models." *Journal of Banking & Finance*, Vol. 24, pp. 119-149.
- Gordy, M. B. and E. Heitfield, 2002, "Estimating Default Correlations from Short Panels of Credit Rating Performance Data."
- Hu, Z., A. Levy, and J. Zhang, 2012, "Economic Capital Model Validation: A Comparative Study." Moody's Analytics White Paper.
- Hu, Z., A. Levy, and J. Zhang, 2010, "Navigating Through Crisis: Validating RiskFrontier® Using Portfolio Selection." Moody's Analytics White Paper.
- Huang, J., N. Patel, and L. Pospisil, 2012, "Modeling Correlations of Small and Medium Enterprises." Moody's Analytics White Paper.
- Levy, A., 2008, "An Overview of Modeling Credit Portfolios." Moody's Analytics White Paper.
- Levy, A. and Z. Hu, 2007, "Incorporating Systematic Risk in Recovery: Theory and Evidence," Moody's Analytics White Paper.
- Levy, A., N. Patel, L. Pospisil, and V. Sesum, 2011, "Modeling Sovereign Correlations." Moody's Analytics White Paper.
- Meng, Q., A. Levy, A. Kaplin, Y. Wang, and Z. Hu, 2010, "Implications of PD-LGD correlation in a Portfolio Setting," Moody's Analytics White Paper.
- Vasicek, O., 2002, "Loan Portfolio Value," *Risk*.
- Zeng, B. and J. Zhang, 2001. "An Empirical Assessment of Asset Correlation Models." Moody's Analytics White Paper.
- Zhang, J., F. Zhu, and J. Lee, 2008, "Asset Correlation, Realized Default Correlation, and Portfolio Credit Risk." Moody's Analytics White Paper.