LSE Department of Statistics Practitioners' Challenge 2024 Report

Gurinderjit Ladhar, Wee Tze Lai, Premya Rajakumar, Jiayi Liu, and Vishal Kumar Lohana

February 28, 2024

1 Introduction

This report aims to delve into the methodologies utilised in credit risk assessment, shedding light on the intricate dynamics of default probabilities and correlations. To achieve this, the report is structured as follows: the significance of credit risk, detailed explanations of the Method of Moments, the James-Stein Estimator, rolling window correlations, and the Merton model. Then, we provide an overview of the methodology employed in this study, and the findings of our analysis are presented, with critical evaluation. Finally, we conclude with a summary of key findings and recommendations for future research in the field of credit risk assessment.

2 Credit Risk

Credit risk is the possibility of a loss resulting from a borrower's failure to repay a loan or meet contractual obligations. Traditionally, it refers to the risk that a lender may not receive the owed principal and interest. Credit risk can be estimated using default probabilities, historical data, use of bond prices, credit default swaps etc. More specifically, throughout this report, we are looking at the possibility of financial losses due to a default event i.e. an event where a firm fails to meet its debt obligations.

We understand the financial condition of firms in the same industry or within the same country may reflect similar factors, and so may improve or deteriorate in a correlated fashion, we aim to show this by showing four correlation coefficients. The four coefficients are:

1. One Global Correlation Coefficient: Average correlation coefficient between the returns of any two companies selected from different countries and different sectors, in the Latin American and Caribbean region.
2. One Country Correlation Coefficient: Average correlation coefficient between the returns of any two companies selected from the same country, but from different industry sectors, in the Latin American and Caribbean region.

3. One Financial Sector Correlation Coefficient: Average correlation coefficient between the returns of any two companies selected in the financial sector and from the same country, in the Latin American and Caribbean region.
4. One Non-Financial Sector Correlation Coefficient: Average correlation coefficient between the returns of any two companies selected in the non-financial industry sector and from the same country, in the Latin American and Caribbean region.

There are two types of models to model credit risk, structural models and reduced-form models. Structural models provide a relation between default risk and capital structure. Within a structural model, a default event is deemed to occur for a firm when its assets reach a sufficiently low level compared to its liabilities. These models require strong assumptions on the dynamics of the firm's asset, its debt and how its capital is structured.
Whereas a reduced form model is one that typically uses historical market data, such as stock prices, interest rates, and default rates, to directly estimate the probability of default credit risk metrics without explicitly modelling the underlying economic or financial variables that drive default. It attempts to explicitly model the underlying factors affecting credit risk, such as firm value, asset volatility, and leverage.

The initial interest in credit risk models stemmed from the desire to develop more rigorous quantitative estimates of the amount of economic capital needed to support a bank's risk-taking activities. But, because of the scarcity of the data stemming from the infrequent nature of default events and the longer-term time horizons used in measuring credit risk, models are not only desired but also necessary

3 Theoretical Framework

There were many statistical methods used to model credit risk. This section provides a summary of the methodologies to establish 1. Method of moments, 2. the James-Stein estimator 3. rolling window correlations, and 4. The Merton model.

3.1 Method of Establishing Method of Moments

Let's indicate with D_t the number of defaults in year t and with N_t the number of exposures at the beginning of year t.

The number of the observed joint (pair) defaults is:

$$\binom{D_t}{2} = \frac{D_t(D_t-1)}{2}$$

The total number of possible joint (pair) defaults is:

$$\binom{N_t}{2} = \frac{N_t(N_t-1)}{2}$$

The joint default rate for year t is:

$$p_{joint,t} = \frac{D_t(D_t - 1)}{N_t(N_t - 1)}$$

and the average joint default rate for all t's is:

$$p_{joint} = \frac{1}{T}\sum_{t=1}^{T} p_{joint,t} = \frac{1}{T}\sum_{t=1}^{T} \frac{D_t(D_t - 1)}{N_t(N_t - 1)}$$

The default correlation is specified by the individual and joint default probabilities. So using the average default and joint default probabilities, we calculate the default correlation as

$$\rho_{i,j} = \frac{p_{i,j} - p_i p_j}{\sqrt{p_i(1 - p_j)p_j(1 - p_i)}}$$

When the asset value Ai drops below a critical value di (called default threshold), default is triggered.

The default indicator for the obligor i can be represented as

$$Default_i = 1\,, if\; A_i <= d_i$$

Whereas the no default indicator is

$$No\,default_i = 0\,, if\; A_i > d_i$$

The joint default probability is instead written as

$$Prob(A_i <= d_i\,, A_j <= d_j)$$

continuing, the asset value Ai depends on the common factor Z, which is common to all obligors, on the idiosyncratic factor, ei and on the factor sensitivity wi

$$A_i = w_i Z + \sqrt{1 - w_i^2}\epsilon_i$$

The asset correlation between i and j is completely determined by their factor sensitivities wi and wj and simplifies to

$$\rho_{i,j}^{asset} = w_i w_j$$

3.2 Methods of Shrinking Matrices

One popular method for shrinking matrices is the James-Stein Estimator. It is defined through the following formula for means.

$$\hat{\theta} = \left(1 - \frac{p-2}{\|X\|^2}\right) X$$

It makes values greater than the grand average smaller, and values smaller than the grand average, greater. By shrinkage, we mean moving the values towards the average, or zero in some cases.

3.3 Methods of Establishing Rolling Windows

(Yeo, 2020) Two popular methods for establishing rolling windows are the rolling window estimator and the exponential smoother. The rolling correlation estimator is defined for returns with zero mean and can be computed through the following formula:

$$\hat{\rho}_{12,t} = \frac{\sum_{s=t-n-1}^{t=1} r_{1,s} \cdot r_{2,s}}{\sqrt{(\sum_{s=t-n-1}^{t=1} r_{1,s}^2) \cdot (\sum_{s=t-n-1}^{t=1} r_{2,s}^2)}}$$

This method provides estimates in the range [−1, 1] and gives equal weight to all observations fewer than n periods in the past and zero weight to observations older than that.

An alternative method, which we were not able to look at, does not pick a specific termination point but emphasises current observations by assigning greater weights to more recent observations is the exponential smoother, which is given by the formula:

$$\hat{\rho}_{12,t} = \frac{\sum_{s=1}^{t=1} \lambda^{t-s-1} \cdot r_{1,s} \cdot r_{2,s}}{\sqrt{(\sum_{s=1}^{t=1} \lambda^{t-s-1} \cdot r_{1,s}^2) \cdot (\sum_{s=1}^{t=1} \lambda^{t-s-1} \cdot r_{2,s}^2)}}$$

3.4 Methods of Establishing the Merton Model

In the Merton (1974) framework, equity is viewed as a European call option on the underlying market value of the firm's assets, with a strike price equal to the face value of its debt.

Assuming the firm's assets value follows the stochastic process:

$$dV_A = \mu V_A dt + \sigma_A V_A dz$$

where VA is the firm's asset value, dVA is the change in asset value, μ and σA is the drift rate and volatility of a firm's asset value, and dz is a standard Wiener process.

The capital structure in this framework allows for a single class of debt and equity. If F is the book value of the debt which is due at time T, then the market value of equity and the market value of assets (using the Black-Scholes formula) are related by the following expression:

$$V_E = V_A N(d_1) - e^{-rT} F N(d_2) \tag{1}$$

where VE is the market value of the firm's equity, VA is the market value of assets, r is the risk-free interest rate, and N(d1) and N(d2) are the standard cumulative normal of d1 and d2 given as:

$$d_1 = \frac{ln(\frac{V_A}{F}) + (r + \frac{\sigma_A^2}{2})T}{\sigma_A \sqrt{T}}$$

$$d_2 = d_1 - \sigma_A \sqrt{T}$$

Using Ito's lemma, the equity volatility (σE) and asset volatility (σA) are related by the expression:

$$\sigma_E = \frac{V_A}{V_E} \cdot \frac{\partial V_E}{\partial V_A} \sigma_A$$

From the Black-Scholes options pricing formula, ∂V_E ∂V_A = N(d1). Thus, the above expression can be rewritten as:

$$\sigma_E = \frac{V_A}{V_E} \cdot N(d_1) \sigma_A \tag{2}$$

In practice, VE, σE, F, T, and r are known. To compute VA and σA, one can solve the system of two nonlinear equations 1 and 2 simultaneously to minimise the sum of the squared errors:

$$e^2 = e_1^2 + e_2^2 \tag{3}$$

where

$$e_1 = V_E - V_A N(d_1) - e^{-rT} F N(d_2)$$

and

$$e_2 = \sigma_E - \frac{V_A}{V_E} \cdot N(d_1)\sigma_A$$

Once the values of VA and σA are obtained, the Distance to default (DtD) is computed as:

$$\mathrm{DtD} = \frac{V_A - F}{V_A \cdot \sigma_A} \tag{4}$$

Then the probability of default is computed as:

$$PD = 1 - N(DD)$$

4 Methodology

4.1 Overview of approaches taken

Having provided an overview of several statistical methods for forecasting conditional correlations, two different approaches were adopted for this project. After obtaining stock price annual log returns, we had one approach to pursue a ‘simple’ average. This meant creating a correlation matrix, where each entry is the correlation of the corresponding symbols for that cell’s stock price annual log returns. Afterwards, we could shrink that matrix, apply filtering for our four desired correlation coefficients, and finally in the filtered matrixes, add up the non-zero cells and divide by the number of non-zero cells. This is an idea of an ‘average over companies’, a simple average as we call it.

The second approach we took was to work out an ‘average over time’, this was done by utilising rolling correlations with a fixed sliding window of size three years. After obtaining the rolling correlations between stock price annual log returns (equity correlation) between all pairs of companies, you could stop there and be satisfied that you can now access the correlation over time between any two companies as the brief required. If not satisfied, the next step was to create multiple matrices, each one corresponding to the number of windows, e.g. the first matrix would contain the first window rolling correlation between all companies. Finally, we averaged all of these matrices, thus obtaining a single matrix containing the averaged rolling correlation for all pairs of companies. From here again, we can shrink the matrix and filter, but this time leave the filtered matrixes as is. This means on average over time, a single correlation coefficient was not attained as adding up the entire correlation over companies leads to a very simplified and exaggerated output. Consider for example if, there were a few negative correlations, but the rest were positive, then overall the negative correlations would not be shown in the final single figure. We then applied the same process for default probability correlations, using the probability of default for each company instead of the stock price annual log returns. We then used the interpretation of default correlation as the ‘correlation in the time series of default probabilities’, as explained by Donald van Deventer. (Deventer, 2015)

4.2 Data preprocessing

As for the historical data, little to no changes were needed as the data was readily available on the S&P website. (S&P, 2021) Before using the stock price data, preliminary processing was important. The stock price dataset to calculate return on equity contained dates ranging from 1997 to 2022. To ensure the integrity and consistency of our results, we eliminated entries with missing values. Consequently, the refined dataset used for analysis and interpretation encloses the years 2008 to 2020. We ensured that all companies chosen fitted the credit rating category needed BBB-C++. Additionally, within our Merton model, one of the crucial inputs involved the debt values (short-term debt + ½ long-term debt) for each company across the years 2008 to 2020. Notably, the debt values presented in the financial statements were denominated in various currencies specific to each company. To ensure uniformity and facilitate comparative analysis, a necessary step involved the conversion of these diverse currency-denominated debt values into a standardised currency, specifically US dollars. This currency conversion process was a crucial component of our data processing, enabling a consistent and comprehensive evaluation of the KMV model results.

Figure 1: Section of the table with stock price data

	symbol	date	open	high	low	close	volume	adjusted
1	VALE	2008-01-02	33.38	33.45	32.45	32.71	12307300	14.80379
2	VALE	2008-01-03	33.04	33.55	32.80	33.09	10542100	14.97577
3	VALE	2008-01-04	32.50	32.71	31.40	31.76	21463700	14.37384
4	VALE	2008-01-07	31.66	32.07	30.10	31.53	28644400	14.26975
5	VALE	2008-01-08	32.07	32.63	30.87	31.23	22818300	14.13398
6	VALE	2008-01-09	31.42	31.80	30.63	31.40	29498300	14.21092

5 Results

Firstly, here is the correlation matrix containing the linear correlation between stock price annual log returns. It is a smaller sample for brevity.

	ABCB	AMX	BVN	CX	EC	ITUB	MELI	PBR	TV	VALE	YPF
ABCB	1.00000000	0.1022018	-0.2778249	0.25211364	-0.09084177	-0.1330496	0.06745789	-0.17836066	0.2676858	-0.15941328	0.28526160
AMX	0.10220176	1.0000000	0.4547169	0.65811292	0.61346562	0.7359155	0.80638036	0.57839591	0.7144979	0.76215381	0.21871031
BVN	-0.27782490	0.4547169	1.0000000	0.33532123	0.69322459	0.7673417	0.30905442	0.74874429	0.1055519	0.77000074	-0.07194860
CX	0.25211364	0.6581129	0.3353212	1.00000000	0.30332579	0.5338038	0.64186400	0.49967900	0.5453330	0.69509597	-0.09236368
EC	-0.09084177	0.6134656	0.6932246	0.30332579	1.00000000	0.6057965	0.29001401	0.60746642	0.2920922	0.70017757	0.15594595
ITUB	-0.13304958	0.7359155	0.7673417	0.53380383	0.60579650	1.0000000	0.52300417	0.86256844	0.4477827	0.83547288	0.31530646
MELI	0.06745789	0.8063804	0.3090544	0.64186400	0.29001401	0.5230042	1.00000000	0.55068250	0.4050199	0.71586051	-0.08171579
PBR	-0.17836066	0.5783959	0.7487443	0.49967900	0.60746642	0.8625684	0.55068250	1.00000000	0.1855767	0.87870630	0.08317181
TV	0.26768583	0.7144979	0.1055519	0.54533303	0.29209217	0.4477827	0.40501985	0.18557673	1.0000000	0.28757178	0.24822225
VALE	-0.15941328	0.7621538	0.7700007	0.69509597	0.70017757	0.8354729	0.71586051	0.87870630	0.2875718	1.00000000	0.05472279
YPF	0.28526160	0.2187103	-0.0719486	-0.09236368	0.15594595	0.3153065	-0.08171579	0.08317181	0.2482223	0.05472279	1.00000000

Figure 2: Correlation matrix of stock price annual log returns

Here is the shrunk correlation matrix, using the James-Stein estimator using a shrinkage intensity of `0.3216. (R, 2023)`

	ABCB	AMX	BVN	CX	EC	ITUB	MELI	PBR	TV	VALE	YPF
ABCB	1.0000000	-0.3734176	-0.6239327	-0.1410490	-0.5724712	-0.62858867	-0.2789519	-0.6142792	0.135102534	-0.5917077	0.218680446
AMX	-0.3734176	1.0000000	0.3251059	0.4877288	0.3526193	0.46019457	0.5891516	0.4042757	0.361308806	0.5043589	-0.319424906
BVN	-0.6239327	0.3251059	1.0000000	0.1962818	0.5984920	0.60855507	0.2638517	0.6331596	-0.207225803	0.6082967	-0.347490162
CX	-0.1410490	0.4877288	0.1962818	1.0000000	0.1076875	0.26602410	0.5583227	0.2920404	0.295476345	0.4038582	-0.503306661
EC	-0.5724712	0.3526193	0.5984920	0.1076875	1.0000000	0.53270245	0.1919816	0.5361536	-0.115642133	0.5352554	-0.232809343
ITUB	-0.6285887	0.4601946	0.6085551	0.2660241	0.5327025	1.00000000	0.3578632	0.6417769	-0.025750183	0.6219160	-0.225734074
MELI	-0.2789519	0.5891516	0.2638517	0.5583227	0.1919816	0.35786322	1.0000000	0.3887315	0.207922056	0.4889739	-0.456541340
PBR	-0.6142792	0.4042757	0.6331596	0.2920404	0.5361536	0.64177692	0.3887315	1.0000000	-0.164809578	0.6524794	-0.340459076
TV	0.1351025	0.3613088	-0.2072258	0.2954763	-0.1156421	-0.02575018	0.2079221	-0.1648096	1.000000000	-0.0655766	-0.007907675
VALE	-0.5917077	0.5043589	0.6082967	0.4038582	0.5352554	0.62191596	0.4889739	0.6524794	-0.065576602	1.0000000	-0.414433116
YPF	0.2186804	-0.3194249	-0.3474902	-0.5033067	-0.2328093	-0.22573407	-0.4565413	-0.3404591	-0.007907675	-0.4144331	1.000000000

Figure 3: Shrunk Correlation matrix, using James-Stein estimator.

Figure 4: Filtered global equity correlation matrix with correlation coefficient `0.139103942012609`

	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11
1	0.0000000	-0.3734176	-0.6239327	-0.1410490	-0.5724712	0.00000000	-0.2789519	0.0000000	0.135102534	0.0000000	0.218680446
2	-0.3734176	0.0000000	0.3251059	0.0000000	0.3526193	0.46019457	0.5891516	0.4042757	0.000000000	0.5043589	-0.319424906
3	-0.6239327	0.3251059	0.0000000	0.1962818	0.5984920	0.60855507	0.2638517	0.6331596	-0.207225803	0.6082967	-0.347490162
4	-0.1410490	0.0000000	0.1962818	0.0000000	0.1076875	0.26602410	0.5583227	0.2920404	0.000000000	0.4038582	-0.503306661
5	-0.5724712	0.3526193	0.5984920	0.1076875	0.0000000	0.53270245	0.1919816	0.0000000	-0.115642133	0.5352554	0.000000000
6	0.0000000	0.4601946	0.6085551	0.2660241	0.5327025	0.00000000	0.3578632	0.0000000	-0.025750183	0.0000000	-0.225734074
7	-0.2789519	0.5891516	0.2638517	0.5583227	0.1919816	0.35786322	0.0000000	0.3887315	0.207922056	0.4889739	0.000000000
8	0.0000000	0.4042757	0.6331596	0.2920404	0.0000000	0.00000000	0.3887315	0.0000000	-0.164809578	0.0000000	0.000000000
9	0.1351025	0.0000000	-0.2072258	0.0000000	-0.1156421	-0.02575018	0.2079221	-0.1648096	0.000000000	-0.0655766	-0.007907675
10	0.0000000	0.5043589	0.6082967	0.4038582	0.5352554	0.00000000	0.4889739	0.0000000	-0.065576602	0.0000000	-0.414433116
11	0.2186804	-0.3194249	-0.3474902	-0.5033067	0.0000000	-0.22573407	0.0000000	0.0000000	-0.007907675	-0.4144331	0.000000000

Figure 5: Filtered country equity correlation matrix with correlation coefficient `0.129606139044811`

	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11
1	0.0000000	0.0000000	0	0.0000000	0	0.0000000	0.0000000	-0.6142792	0.0000000	-0.5917077	0.0000000
2	0.0000000	0.0000000	0	0.4877288	0	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000
3	0.0000000	0.0000000	0	0.0000000	0	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000
4	0.0000000	0.4877288	0	0.0000000	0	0.0000000	0.0000000	0.0000000	0.2954763	0.0000000	0.0000000
5	0.0000000	0.0000000	0	0.0000000	0	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000
6	0.0000000	0.0000000	0	0.0000000	0	0.0000000	0.0000000	0.6417769	0.0000000	0.6219160	0.0000000
7	0.0000000	0.0000000	0	0.0000000	0	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	-0.4565413
8	-0.6142792	0.0000000	0	0.0000000	0	0.6417769	0.0000000	0.0000000	0.0000000	0.6524794	0.0000000
9	0.0000000	0.0000000	0	0.2954763	0	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000
10	-0.5917077	0.0000000	0	0.0000000	0	0.6219160	0.0000000	0.6524794	0.0000000	0.0000000	0.0000000
11	0.0000000	0.0000000	0	0.0000000	0	0.0000000	-0.4565413	0.0000000	0.0000000	0.0000000	0.0000000

Figure 6: Filtered financial equity correlation matrix with correlation coefficient

-0.628588668171016

	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11
1	0.0000000	0	0	0	0	-0.6285887	0	0	0	0	0
2	0.0000000	0	0	0	0	0.0000000	0	0	0	0	0
3	0.0000000	0	0	0	0	0.0000000	0	0	0	0	0
4	0.0000000	0	0	0	0	0.0000000	0	0	0	0	0
5	0.0000000	0	0	0	0	0.0000000	0	0	0	0	0
6	-0.6285887	0	0	0	0	0.0000000	0	0	0	0	0
7	0.0000000	0	0	0	0	0.0000000	0	0	0	0	0
8	0.0000000	0	0	0	0	0.0000000	0	0	0	0	0
9	0.0000000	0	0	0	0	0.0000000	0	0	0	0	0
10	0.0000000	0	0	0	0	0.0000000	0	0	0	0	0
11	0.0000000	0	0	0	0	0.0000000	0	0	0	0	0

Figure 7: Filtered Non-Financial equity correlation matrix with correlation coefficient

0.268090386023197

	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11
1	0	0.0000000	0	0.0000000	0	0	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000
2	0	0.0000000	0	0.4877288	0	0	0.0000000	0.0000000	0.3613088	0.0000000	0.0000000
3	0	0.0000000	0	0.0000000	0	0	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000
4	0	0.4877288	0	0.0000000	0	0	0.0000000	0.0000000	0.2954763	0.0000000	0.0000000
5	0	0.0000000	0	0.0000000	0	0	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000
6	0	0.0000000	0	0.0000000	0	0	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000
7	0	0.0000000	0	0.0000000	0	0	0.0000000	0.0000000	0.0000000	0.0000000	-0.4565413
8	0	0.0000000	0	0.0000000	0	0	0.0000000	0.0000000	0.0000000	0.6524794	0.0000000
9	0	0.3613088	0	0.2954763	0	0	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000
10	0	0.0000000	0	0.0000000	0	0	0.0000000	0.6524794	0.0000000	0.0000000	0.0000000
11	0	0.0000000	0	0.0000000	0	0	-0.4565413	0.0000000	0.0000000	0.0000000	0.0000000

Figure 8: One example of the rolling correlations

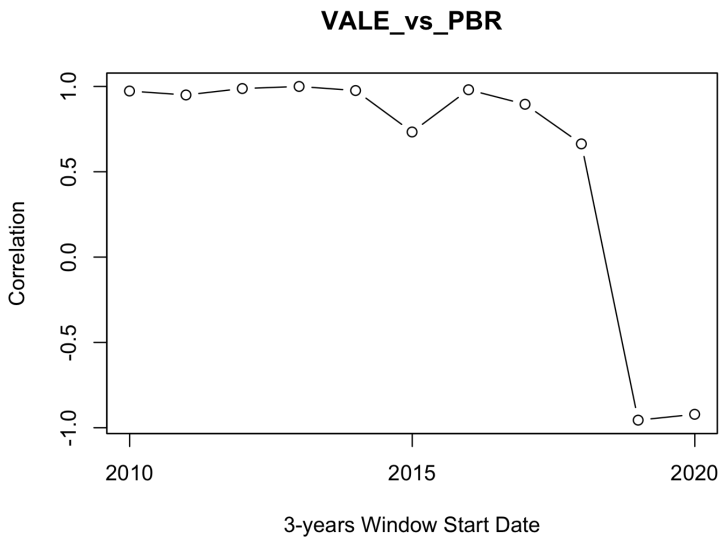


Figure 9: A matrix containing the first window of the rolling correlations

	VALE	PBR	ITUB	ABCB	EC	AMX	CX	BVN	TV	YPF	MELI
VALE	NA	0.9732398	0.9857243	-0.08604847	0.7738490	0.98399366	0.9962794072	0.9676128	0.9823325	-0.8267180	0.99929373
PBR	NA	NA	0.9980356	-0.31268533	0.6075974	0.91671209	0.9498148430	0.8837109	0.9130410	-0.9338796	0.96391750
ITUB	NA	NA	NA	-0.25256316	0.6561628	0.93994276	0.9675465435	0.9112969	0.9368001	-0.9096424	0.97870132
ABCB	NA	NA	NA	NA	0.5644324	0.09287136	0.0001340608	0.1682410	0.1019215	0.6316675	-0.04854991
EC	NA	NA	NA	NA	NA	0.87433129	0.8255549712	0.9086735	0.8787084	-0.2834105	0.79710273
AMX	NA	NA	NA	NA	NA	NA	0.9956905571	0.9971104	0.9999587	-0.7132251	0.98999507
CX	NA	NA	NA	NA	NA	NA	NA	0.9857684	0.9948061	-0.7751547	0.99881424
BVN	NA	NA	NA	NA	NA	NA	NA	NA	0.9977599	-0.6579167	0.97641537
TV	NA	NA	NA	NA	NA	NA	NA	NA	NA	-0.7068218	0.98867106
YPF	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	-0.80499260
MELI	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Figure 10: The single matrix obtained after averaging the rolling correlations

	VALE	PBR	ITUB	ABCB	EC	AMX	CX	BVN	TV	YPF	MELI
VALE	1.0000000	0.57121100	0.57587615	0.13340029	0.6168213	0.6502402	0.9116730	0.53456875	0.4769941	0.22185199	0.42576849
PBR	0.5712110	1.00000000	0.47544413	0.13800079	0.4910914	0.3868651	0.5269766	0.43228222	0.3028470	0.08576818	0.06773288
ITUB	0.5758762	0.47544413	1.00000000	0.06208236	0.4490902	0.5483730	0.5325883	0.68616026	0.4231361	0.43839442	0.44617918
ABCB	0.1334003	0.13800079	0.06208236	1.00000000	0.2852889	0.2439952	0.2282383	-0.23261452	0.1896869	0.43521292	0.19249830
EC	0.6168213	0.49109136	0.44909019	0.28528895	1.0000000	0.7121894	0.5657001	0.52339324	0.6004059	0.48880370	0.30904830
AMX	0.6502402	0.38686511	0.54837299	0.24399516	0.7121894	1.0000000	0.7098188	0.27176760	0.8801363	0.51288626	0.62200347
CX	0.9116730	0.52697662	0.53258833	0.22823832	0.5657001	0.7098188	1.0000000	0.38566576	0.5908473	0.21820157	0.44738189
BVN	0.5345687	0.43228222	0.68616026	-0.23261452	0.5233932	0.2717676	0.3856658	1.00000000	0.1947679	0.09466937	0.05511054
TV	0.4769941	0.30284702	0.42313611	0.18968691	0.6004059	0.8801363	0.5908473	0.19476794	1.0000000	0.28139193	0.39598295
YPF	0.2218520	0.08576818	0.43839442	0.43521292	0.4888037	0.5128863	0.2182016	0.09466937	0.2813919	1.00000000	0.61278827
MELI	0.4257685	0.06773288	0.44617918	0.19249830	0.3090483	0.6220035	0.4473819	0.05511054	0.3959830	0.61278827	1.00000000

The matrix above after being shrunk, (but this is not shown to avoid repetition) and filtered to generate the filtered average over time correlation matrix is shown below.

Figure 11: Average over time filtered global correlation

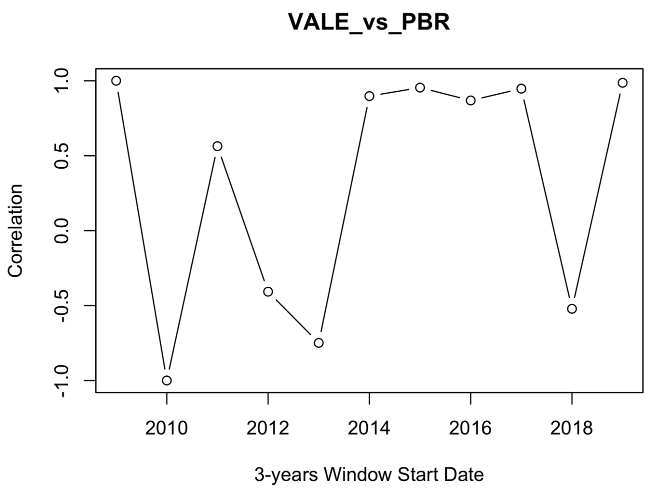
	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11
1	0.00000000	0.00000000	0.00000000	0.00000000	0.26809322	0.24870727	0.53697025	0.31431463	0.20516246	-0.30232719	-0.01175825
2	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	-0.02564941	0.27302891	0.31416751	0.01204259	0.00000000	-0.33086332
3	0.00000000	0.00000000	0.00000000	0.00000000	0.08509802	0.08482411	0.17794576	0.44429506	0.04087321	-0.09903892	0.02762652
4	0.00000000	0.00000000	0.00000000	0.00000000	-0.14919480	-0.09357170	-0.20377551	-0.48458382	-0.08050767	0.23997528	0.02589456
5	0.26809322	0.00000000	0.08509802	-0.14919480	0.00000000	0.31470073	0.23964028	0.22089836	0.30665493	0.00000000	-0.08680422
6	0.24870727	-0.02564941	0.08482411	-0.09357170	0.31470073	0.00000000	0.00000000	-0.04783562	0.00000000	0.08932227	0.29743034
7	0.53697025	0.27302891	0.17794576	-0.20377551	0.23964028	0.00000000	0.00000000	0.17361917	0.00000000	-0.26667385	0.05207669
8	0.31431463	0.31416751	0.44429506	-0.48458382	0.22089836	-0.04783562	0.17361917	0.00000000	-0.03998309	-0.29562272	-0.23293350
9	0.20516246	0.01204259	0.04087321	-0.08050767	0.30665493	0.00000000	0.00000000	-0.03998309	0.00000000	-0.02597760	0.15954132
10	-0.30232719	0.00000000	-0.09903892	0.23997528	0.00000000	0.08932227	-0.26667385	-0.29562272	-0.02597760	0.00000000	0.00000000
11	-0.01175825	-0.33086332	0.02762652	0.02589456	-0.08680422	0.29743034	0.05207669	-0.23293350	0.15954132	0.00000000	0.00000000

Now the same procedure can be repeated for the default probability correlations, but only the graph and one filtered matrix are shown here; the full matrices can be accessed within our GitHub. (Lohana, 2024)

Figure 12: Average over time filtered global default correlation

	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11
1	0.0000000	0.0000000	0.00000000	0.00000000	-0.28983258	0.13689415	0.13628259	0.2193792	-0.22018011	-0.1659816	-0.40256037
2	0.0000000	0.0000000	0.00000000	0.00000000	0.00000000	0.30850191	0.31256006	0.5505181	-0.45202800	0.0000000	-0.32593201
3	0.0000000	0.0000000	0.00000000	0.00000000	0.22011724	0.44805138	0.45128966	0.4028536	-0.04276127	-0.5105491	-0.11571786
4	0.0000000	0.0000000	0.00000000	0.00000000	0.09178700	0.28980866	0.11233273	-0.1987606	0.09963834	-0.1171002	-0.16996939
5	-0.2898326	0.0000000	0.22011724	0.09178700	0.00000000	-0.01631037	-0.28470980	-0.1811591	0.15525883	0.0000000	0.12722075
6	0.1368941	0.3085019	0.44805138	0.28980866	-0.01631037	0.00000000	0.00000000	0.4295222	0.00000000	-0.6343081	-0.49682634
7	0.1362826	0.3125601	0.45128966	0.11233273	-0.28470980	0.00000000	0.00000000	0.4110914	0.00000000	-0.4490415	-0.05537106
8	0.2193792	0.5505181	0.40285358	-0.19876062	-0.18115910	0.42952215	0.41109139	0.0000000	-0.41086814	-0.5583508	-0.35982233
9	-0.2201801	-0.4520280	-0.04276127	0.09963834	0.15525883	0.00000000	0.00000000	-0.4108681	0.00000000	0.3099913	0.52939531
10	-0.1659816	0.0000000	-0.51054910	-0.11710022	0.00000000	-0.63430806	-0.44904149	-0.5583508	0.30999129	0.0000000	0.00000000
11	-0.4025604	-0.3259320	-0.11571786	-0.16996939	0.12722075	-0.49682634	-0.05537106	-0.3598223	0.52939531	0.0000000	0.00000000

Figure 13: A plot showing the rolling default probability correlation of two companies



Overall, from our results, equity correlation provides, at best, a very noisy indicator of default correlations. This is not surprising as equity returns incorporate a lot of noise which are not related to firms' fundamentals.

6 Evaluation

6.1 Limitations of Method of Moments

The main limitation for this comes from the two main assumptions which are:
1) ‘All obligors share the same default and joint default probabilities as the other obligors in the same credit group, either Investment Grade or Speculative Grade’
2) ‘There is no serial correlation in the time series of defaults, so defaults in any year are not affected by the defaults in a previous year.’
What this means is that there is not enough data/sufficient data concerning the default correlation among and between specific industries and defaultable bonds, it does not use firm-specific information. Finally, credit risk models should be based on current, rather than historical measurements.

6.2 Limitations of using James-Stein Estimator

We used a constant ‘w’ term in the cor.shrink function when estimating the shrinkage factor, lambda. A constant ‘w’ means uniform weights are assumed (w = rep(1/n, n) with n = nrow(x)). Also, the positive James-Stein estimator is an even better tool as with the regular James-Stein estimator the estimated shrinkage factors can be negative.

6.3 Limitations of using the Merton Model

The main limitation of the Merton model is again the assumptions used throughout such as
1) Assuming that the PD follows a normal distribution which is not the best assumption if we would like to calculate the true probability of default since the actual distribution likely has much ‘fatter’ tails.
2) We assumed a fixed risk-free rate, it can be easily added to our model’s calculations but we simply were not able to collect consistent information on the risk-free rate given the time available, so we set it to be fixed.
3) Assuming $\mu=1$, as μV is difficult to estimate and sensitive to errors.
4) An issue we recognise is the issue of conditional correlation ie in our matrix the probability of default for one company correlated with another company might appear to be strongly correlated however, this may be due to another factor (the ‘k’), we would need to calculate the partial pairwise correlations. Perhaps a better model of default correlation would be to look at the other interpretations mentioned by Donald van Deventer such as looking for default drivers as default correlation. But there is still a link between default probabilities and default correlation, which can be read more in the paper ‘Default probabilities and default correlations’. To briefly explain,

$$1\{Z_1 \leq z_1\} \text{ and } 1\{Z_2 \leq z_2\}.$$

Z1 and Z2 are return variables where z1, and z2 are default points such that if the returns fall below, i.e. Z1<z1 or Z2<z2 then it’s a default. A higher default probability means z1,z2 shift to the right so any shift of z1, z2 is the same as saying higher probability. The first main result is as below:

(i) *If* $p_1, p_2 < 50\%$, *then default correlations* **increase** $(\rho_1^{def} + \rho_2^{def} > 0)$;

(ii) *If* $p_1, p_2 > 50\%$, *then default correlations* **decrease** $(\rho_1^{def} + \rho_2^{def} < 0)$.

If z(i) < 0 (i = 1,2), then a homogeneous shift of z(1) and z(2) to the right leads to reduced skewness of the binary variables, 1{Z1 < z1} and 1{Z2 < z2}. Reduced skew implies more information is revealed about correlated underlying variables Z1 and Z2 thus default correlations increase towards the higher correlations of returns. If z1, z2 > 0, skewness would increase as they move right hence decreasing the default correlation.

The second main result is as follows:

(i) *If* z_1 *moves, then default correlations* **increase** $(\rho_1^{def} > 0)$ *if* $\rho < \rho_+(\lambda)$. *Note that this inequality is fulfilled if* $\lambda \leq 96\%$ *or if* $\rho \leq 56\%$.

(ii) *If* z_2 *moves, then default correlations*

- **increase** $(\rho_2^{def} > 0)$ *if* $\rho < \rho_+(1/\lambda)$
- **decrease** $(\rho_2^{def} < 0)$ *if* $\rho > 2\lambda/(1+\lambda^2)$.

$$\rho_+(\lambda) := (25/32)\left\{\lambda - \sqrt{\lambda^2 + 24/25}\right\}$$

Given λ=z1/z2, If the smaller point, z1, (as z1<z2 is assumed) moves right then as before, skewness falls thus the default correlation increases. And if z2 moves right then the effect depends on the value of the correlation. (Deutsche Bank, 2001)

6.4 Alternative models

There are many alternative models such as the ZPP model, Zhou model, Black Cox model, LT model, etc. All use more complex models such as first-time passage models or copula functions, for which we were unable to model given our limited knowledge and the timeframe. However, we appreciate that alternative views of default correlation may have served better such as default correlation being defined as the correlation of Bernoulli distributed random variables as in the Zhou model.

7 Conclusion

In conclusion, this report delved into the intricate landscape of credit risk assessment, focusing on methodologies like structural and reduced-form models. By exploring methods such as the Method of Moments, the James-Stein Estimator, rolling window correlations, and the Merton model, we aimed to shed light on the complex dynamics underlying credit risk estimation. Our analysis highlighted the challenges and limitations inherent in each approach, emphasising the need for robust models capable of capturing the nuanced interplay of factors influencing credit risk.

References

Deutsche Bank. (2001). *Default probabilities and default correlations*. ECONSTOR.

https://www.econstor.eu/bitstream/10419/40256/1/341385050.pdf

Deventer, D. v. (2015, September 24). *Modeling 3 Meanings Of Correlated Default: A Worked Example*. Seeking Alpha. Retrieved February 28, 2024, from https://seekingalpha.com/article/3532446-modeling-3-meanings-of-correlated-default-worked-example

Lohana, V. (2024, February 14). *LSE Practioner's Challenge 2024*. GitHub. Retrieved February 28, 2024, from https://github.com/vishalk4u/LSE-Practitioner-Challenge-2024/tree/main

R. (2023, December 11). *How to Use the corpcor Package in R*. R Basics. Retrieved February 28, 2024, from https://rbasics.org/guides/how-to-use-the-corpcor-package-in-r/

S&P. (2021, September 13). *Default, Transition, and Recovery: 2020 Annual Latin American Corporate Default And Rating Transition Study*. S&P Global. Retrieved February 28, 2024, from https://www.spglobal.com/ratings/en/research/articles/210913-default-transition-and-recovery-2020-annual-latin-american-corporate-default-and-rating-transition-study-12064230

Yeo, H. (2020, March 11). *LSE Practioner's Challenge 2020*. GitHub. Retrieved February 28, 2024, from https://github.com/huddyyeo/LSE_practitioners_challenge/blob/master/Writeups/Final_Report.pdf