A modelling framework for Expected Credit Loss calculation and Stress Testing for low-default portfolios

Global Credit Data - North American Conference

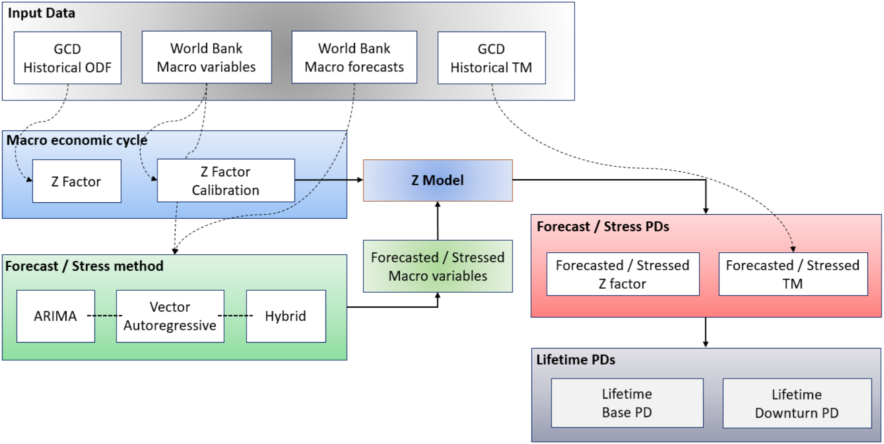
Fausto Molinari

New York, October 1-2 2019

Some disclaimer

All views expressed in this presentation are my own and do not represent the opinions of any entity whatsoever with which I have been, am now, or will be affilated.

Modelling phases - flow



Asymptotic Single Risk Factor framework I

$$\Delta \ln A_{i,t} = -\sqrt{\rho} Z_t + \sqrt{1-\rho}\epsilon_{i,t} + g_i \tag{1}$$

Where:

- $\Delta \ln A_{i,t}$ = company's i log-assets value
- ρ = correlation coefficient between A and Z
- Z_t = systemic risk factor, economic cycle effect
- $\epsilon_{i,t}$ = idiosyncratic risk factor

Per definition:

- $Z \sim \mathcal{N}(0,1)$
- $\epsilon \sim \mathcal{N}(0,1)$
- $COV(\epsilon_i, \epsilon_j) = 0$
- $COV(\epsilon_i, Z) = 0$

Asymptotic Single Risk Factor framework II

Default probabilities and rating migrations can be represented by the probability that a *standard normal* variable falls below a certain critical value.
Asymptotic Single Risk Factor - PD:

$$PD_i | \left[Z_t = \Phi^{-1}(\alpha) \right] = \Phi \left[\frac{\Phi^{-1} \bar{p}_i + \sqrt{\rho} \Phi^{-1}(\alpha)}{\sqrt{1-\rho}} \right] = \Phi \left[\frac{\Phi^{-1} \bar{p}_i - \sqrt{\rho} Z_t}{\sqrt{1-\rho}} \right] \tag{2}$$

Where:

- $PD_i | \left[Z_t = \Phi^{-1}(\alpha) \right]$ = conditional Default Probability given Z_t
- Φ = cumulative normal distribution function
- $\bar{p}_i$ = long run Default Probability for rating i
- α = probability that the Z_t-value (or lower) occurs.

The Z Factor I

Z-series formula derived from the ASRF-PD:

$$Z_t = \frac{\Phi^{-1}\bar{p}_i + \Phi^{-1}PD_i\sqrt{1-\rho}}{\sqrt{\rho}} \tag{3}$$

The model is estimated using the **Method of Moments**
Define m and σ as the average and standard deviation of the series $\Phi^{-1}PD_i$:

$$m = \mathbb{E}\left[\Phi^{-1}PD_i\right] = \frac{\Phi^{-1}\bar{p}_i}{\sqrt{1-\rho}} \tag{4}$$

$$\sigma = \mathbb{V}\left[\Phi^{-1}PD_i\right] = \frac{\rho}{1-\rho} \tag{5}$$

Calculate 4 and 5, solve for $\Phi^{-1}\bar{p}_i$ and ρ:

$$Z_t = \frac{m - \Phi^{-1}PD_i}{\sigma} \tag{6}$$

The Z Factor II

The Z_t series is a *standard normal variable* that describes the economic cycle at time t.

The Z factor will be used to:

- Generate a macroeconomics cycle proxy
- Generate a Point-In-Time transition matrix from the historical (Through-The-Cycle) transition matrix.
- Generate a stressed transition matrix

The Z Factor III

Z Factor calculation:

1. Take a time series of default frequencies (e.g. annual default frequencies for large corporate Europe)
2. Calculate m from 4
3. Calculate σ from 5
4. Obtain ρ from 4 and 5
5. Calculate Z_t by applying 6

Z factor calibration I

The Z series computed according to formula 6 is calibrated using a set of macroeconomics indicators.

$$Z_t = \alpha + \beta_1 Z_{t-1} + \beta_m M_{t,m} + \epsilon_t \tag{7}$$

Where:

- α and β: *intercept* and *slope* of the regression model, estimated via *Generalised Least Square* method
- Z_{t-1}: one-period lag of the Z series
- $M_{t,m}$: t x m matrix of macroeconomics variables
- ϵ_t: regression error term, $\epsilon_t \sim \mathcal{N}(0, \sigma)$

The regression model generates a new Z-series adjusted to the macroeconomics cycle.

Z factor calibration II

Z-modelling steps:

1. Pre-modelling
   1.1 Consistency of the input variables (scale, outliers, missing values)
   1.2 Multi correlation analysis
   1.3 Stationarity of regressors
   1.4 Combination and/or transformations (lag, difference, rate)
2. Modelling
   2.1 Estimate the regression model 7
   2.2 Significance of the model as a whole (F-test)
   2.3 Significance of single regressors (t-test)
   2.4 Explanatory power (R^2)
   2.5 Coefficients signs!
3. Post-modelling
   3.1 Residual analysis (autocorrelation, heteroschedasticity)
   3.2 Model (coefficients) stability

Z factor forecast and stress I

The regression parameters estimated by model 7 are exploited to forecast and stress the Z-series:

$$\begin{aligned} \textit{Normal Forecast} \quad &\rightarrow \quad Z_{t+\tau} = \widehat{\alpha} + \widehat{\beta_1} Z_{t+\tau-1} + \widehat{\beta}_m M_{t+\tau,m} \quad (8) \\ \textit{Stressed Forecast} \quad &\rightarrow \quad \widetilde{Z}_{t+\tau} = \widehat{\alpha} + \widehat{\beta_1} Z_{t+\tau-1} + \widehat{\beta}_m \widetilde{M}_{t+\tau,m} \quad (9) \end{aligned}$$

Where:

- τ: forecast horizon, 1, 2, ..., T
- $\widehat{\alpha}$ and $\widehat{\beta}$: estimated regression parameters
- $Z_{t+\tau-1}$: one-period lag of the forecasted Z-series
- $M_{t+\tau,m}$: $t \times m$ matrix of macroeconomics variables forecasts
- $\widetilde{M}_{t+\tau,m}$: $t \times m$ matrix of stressed macroeconomics variables

Z factor forecast and stress II

In order to apply model 8 or 9, the input matrix M or $\widetilde{M}$ must contain a consistent set of forecasts for all the input variables included in model 7.

How $M/\widetilde{M}$ can be created:

- External source
  - World Bank (Global Economic Prospects)
  - National Bureaus of Statistics
- Quantitative method
  - ARIMA model
  - VAR model
- Hybrid
- Expert judgement

ARIMA model

An Auto Regressive Integrated Moving Average model is:

$$ARIMA(p, d, q) = \pi(L)^p(1 - L)^d y_t = \pi_0 + \theta(L)^q \epsilon_t \tag{10}$$

Where:

- y_t: univariate time series
- $\pi(L)^p$: AR(p) operator
- $\theta(L)^q$: MA(q) operator
- $(1 - L)^d$: Integration (d) operator
- ϵ_t: non observable stationary white-noise process.

Example ARIMA(1,0,1):

$$y_t = \pi_0 + \pi_1 y_{t-1} + \epsilon_t + \theta\epsilon_{t-1} \tag{11}$$

VAR model I

A vector autoregressive model of order p, VAR(p), is a multivariate autoregressive model or as a system of AR(p) models .

General VAR(p) model formulation:

$$VAR(P) \rightarrow \Pi(L)Y_t = c + \Pi_1 Y_{t-1} + \ldots + \Pi_p Y_{t-p} + \epsilon_t \qquad (12)$$

Where:

- Y: (t x 1) vector of time series
- Π_p: (m x m) coefficient matrix
- c: (t x 1) vector of constants
- ϵ: (t x 1) non observable stationary *white noise* process.

VAR model II

Given a VAR(1) model, the maximum likelihood estimation of the unrestricted reduced form is OLS:

$$\widehat{\Pi}(1) = \left(\sum_{t=1}^{T} Y_{t-1} Y_{t-1} \right)^{-1} \left(\sum_{t=1}^{T} Y_t Y_{t-1} \right) \tag{13}$$

While the Mean Squared Error matrix of the k-step forecast is:

$$MSE(k) = \sum_{y}(k) = \sum_{j=0}^{k-1} \Phi_j \Sigma_e \Phi_j^{'} \tag{14}$$

Matrix $\widehat{\Pi}$ contains the estimated VAR coefficients that will be used to produce multivariate forecasts and stress of the macro variables while *MSE* will be used to build the coefficient intervals.

VAR model III

Example:
bivariate VAR(1) for GDP Growth (g) and Unemployment (u) given the estimated coefficient matrix $\widehat{\Pi}$.

$$VAR(1) \rightarrow \begin{pmatrix} g_t \\ u_t \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} + \begin{pmatrix} \hat{\pi}_{11} & \hat{\pi}_{12} \\ \hat{\pi}_{21} & \hat{\pi}_{22} \end{pmatrix} \begin{pmatrix} g_{t-1} \\ u_{t-1} \end{pmatrix}$$

Or:

$$VAR(1) = \begin{cases} g_t = c_1 + \hat{\pi}_{11} g_{t-1} + \hat{\pi}_{12} u_{t-1} \\ u_t = c_2 + \hat{\pi}_{12} g_{t-1} + \hat{\pi}_{22} u_{t-1} \end{cases}$$

VAR model IV

Z factor forecast procedure:

1. Take the macro variables used to calibrate the Z factor in 7
2. Generate at least one forecast for these macro variables
   - External Source, ARIMA: if M is univariate
   - External source + VAR: if M is multivariate
3. Take the regression parameters α and β estimated in 7 and implement equation 9 using the forecasted M.
4. For each forecasted future value of M a new Z factor value is generated.

Scenario definition and stress I

IFRS9: normal-low business cycle $\rightarrow$ 1 over 7 years
Stress testing: 1 over 25 years

Univariate historical shift:

$$\delta_m = \text{MAD}(M_m) \text{ or } \text{SD}(M_m) \tag{15}$$

$$\widetilde{M}_{t+\tau,m} = M_{t+\tau,m} * (1 + \delta_m) \tag{16}$$

Univariate historical p-quantile:

$$\widetilde{M}_{t+\tau,m} = Q_p(M_m|_1^{t+\tau-1}), \quad p \approx 0.14 \tag{17}$$

Multivariate case - Historical shift/quantile & VAR model:

1. Take a *relevant* variable and compute a stressed forecast using 15+16 or 17

Scenario definition and stress II

2. Exploit the new stressed variable together with the estimated VAR parameters in order to compute forecasts for the remaining variables

$$\widetilde{M}_{t,m+\tau} = \begin{cases} M_{t+\tau,m} * (1+\delta_m) & ,\forall \text{ m} = \text{i} \\ \widehat{\Pi}_p M_{t+\tau-p,m} & ,\forall \text{ m} \neq \text{i} \end{cases} \quad (18)$$

Multivariate case - VAR model simulation:

1. Take VAR model estimated in 13
2. Produce a k-step ahead forecast applying a Gaussian shock to the forecast error covariance matrix 14
3. Build a confidence interval for the forecasted value based on the stressed error covariance matrix
4. Take the lower confidence interval as stressed forecast value of the corresponding macro variable.

Scenario definition and stress III

Z stress procedure:

1. Take the matrix $M_{t+\tau,m}$ containing the forecasted macro variables
2. If $M_{t+\tau,m}$ is univariate compute stressed using either historical shift (15 16), quantile 17 or ARIMA simulation.
3. If $M_{t+\tau,m}$ is multivariate combine historical shift method with VAR or run VAR simulation
4. Take the regression parameters α and β estimated in 7 and implement equation 9 using the stressed forecasted $\widetilde{M}_{t+\tau,m}$.
5. For each stressed forecasted future value of $\widetilde{M}$ a new (stressed) Z factor value is generated.

Point-in-Time adjustment

The calibrated/forecasts/stressed Z factor is the key driver used to transform a Ttc transition matrix into Pit.

$$P^{Pit}_{r,c} = \Phi\left[\frac{\Phi^{-1}\left(P^{Ttc}_{r,D} + \ldots + P^{Ttc}_{r,c}\right) - \sqrt{\rho}\widehat{Z}_t}{\sqrt{1-\rho}}\right] - P^{Pit}_{r,D} - \ldots - P^{Pit}_{r,c+1} \tag{19}$$

Where:

- $r, c \in [1, D]$
- $P_{r,c}$, transition probability from rating **r** to **c**
- $\widehat{Z}_t$, forecasted / stressed Z-series
- ρ, Z-series ρ computed from 4 and 5
- Φ, Normal cumulative distribution function
- Φ^{-1}, Inverse Normal distribution function

Lifetime Probability of Default I

Markov Generator Matrix
Given a transition matrix **P** (n x n) that satisfies the following conditions:

- **P** is a matrix with row sum equal to 1
- **P** has no negative entries
- **P** is strictly diagonal dominant, diagonal entries all greater than 0.5

the series:

$$\widetilde{Q} = (P - I) - \frac{(P-I)^2}{2} + \frac{(P-I)^3}{3} + \ldots \tag{20}$$

constitutes a unique generator matrix for **P**. The series converges geometrically quick, and gives rise to an n x n matrix having row sum 0, such that:

$$\exp(\widetilde{Q}) = P.$$

Lifetime Probability of Default II

Correction for negative entries in $\widetilde{Q}$:
Replace the negative values with 0, and add the appropriate value back into the corresponding diagonal entry to preserve the property of row sums 0.

$$G_i = |\widetilde{q_{ii}}| + \sum_{i \neq j} max(\widetilde{q_{ij}}, 0) \tag{21}$$

$$B_i = \sum_{i \neq j} max(-\widetilde{q_{ij}}, 0) \tag{22}$$

$$q_{ij} = \begin{cases} 0, & i \neq j \text{ and } \widetilde{q_{ij}} < 0 \\ \widetilde{q_{ij}} - \frac{B_i |\widetilde{q_{ij}}|}{G_i}, & G_i > 0 \\ \widetilde{q_{ij}}, & G_i = 0 \end{cases} \tag{23}$$

Lifetime Probability of Default III

The generator matrix created according to 20 is used to generate the τ-year Pit matrices beyond the available $Z_{t+\tau}$ values.
The mathematical function used to transform $\widetilde{Q}$ to P(τ) is the **matrix exponential**:

$$P^{Pit}(\tau) = \exp(\tau\widetilde{Q}) = I + \tau\widetilde{Q} + \frac{\left(\tau\widetilde{Q}\right)^2}{2!} + \frac{\left(\tau\widetilde{Q}\right)^3}{3!} + \dots \qquad (24)$$

The cumulative Pit matrix up to time τ is obtained by successive multiplication:

$$\bar{P}^{Pit}(\tau) = \prod_{t=1}^{\tau} P^{Pit}(t) \qquad (25)$$

Lifetime Probability of Default IV

Mean reverting Z:

Alternatively to the generator matrix method, the lifetime PDs can be obtained by assessing the unknown future values of the Z series ($Z_{t+\tau+\lambda}$).

One approach consists of assuming that the lifetime Z is **mean reverting** toward the long-run Z average:

$$Z_{t+\tau+\lambda} = \begin{cases} (Z_\tau) * \left(\frac{\bar{Z}_{LR}}{Z_\tau}\right)^{\lambda/LR} \\ (Z_\tau + 1) * \left(\frac{\bar{Z}_{LR}+1}{Z_\tau+1}\right)^{\lambda/LR} & \text{, if } (Z_\tau * \bar{Z}_{LR}) < 0 \end{cases} \tag{26}$$

Where:

$\bar{Z}_{LR}$ = long-run Z average

Lifetime Probability of Default V

Lifetime PD procedure:

1. Take the forecasted Z values obtained in 8
2. Apply model 19 using Z value today in order to obtain today's Pit TM from Ttc TM
3. Apply model 19 using available forecasted Z values to obtain future Pit TMs
4. Apply 20 + 24 to generate future Pit TMs for the remaining lifetime where Z forecasts are not available
5. Alternatively, use 26 to generate a series of Z values (from last available forecast to lifetime) mean reverting to the long-run average Z.
6. Extract last column from each Pit matrix in order to obtain the final PDs
7. *Repeat Steps 1-6 using the stressed Z in order to obtain the stressed PDs*

Thank you!

fausto.molinari@runbox.com

https://onexcent.com/banking-and-finance/

References I

- 2012, Petrov A., Carlehed M., *A methodology for point-in-time through-the-cycle probability of default composition in risk classification systems*, Journal of Risk Model Validation, Volume 6 Number 3, Fall 2012 (3-25).
- 2001, Israel R., Rosenthal J., Wei, J., *Finding Generator For Markov Chains Via Empirical Transition Matrices, With Applications To Credit Ratings, Mathematical and Finance*, Vol. 11, No. 2, April 2001, 245-265.
- 2014, Yang B., *Modeling Systematic Risk and Point-in-Time Probability of Default under the Vasicek Asymptotic Single Risk Factor Model Framework*, MPRA Paper No. 59025, `http://mpra.ub.uni-muenchen.de/59025/`.
- 2003, Gordy M. B., *A risk-factor model foundation for ratings-based bank capital rules*, Journal of Financial Intermediation 12, 199 - 232.
- 1974, Merton R. C., *On the pricing of corporate debt: The risk structure of interest rates*, Journal of Finance 29, 449 - 470.

References II

- 2002, Vasicek O., *Loan portfolio value*. RISK, December 2002, 160 – 162
- 2009, Kupiec P. H., *How Well Does the Vasicek-Basel AIRB Model Fit the Data? Evidence from a Long Time Series of Corporate Credit Ratings Data*, Working Paper, FDIC Center for Financial Research.
- 1995, Johansen S., Likelihood-Based Inference in Cointegrated Vector Autoregressive Models, Oxford: Clarendon Press.
- 1994, Hamilton J. D., *Time Series Analysis*, Princeton University Press, Princeton, New Jersey.
- 2007, Luetkepohl H., *Econometric Analysis with Vector Autoregressive Models*, Economics working Papers ECO2007/11, European University Institute, `https://ideas.repec.org/p/eui/euiwps/eco2007-11.html`.
- 1990, Harvey A. C., *Forecasting, structural time series models and the Kalman filter*, Cambridge University Press.

References III

- 1962, Zellner A., *An efficient method of estimating seemingly unrelated regression equations and tests for aggregation bias*, Journal of the American Statistical Association. 57: 348–368.