

Université Paris 13

Sup' Galilée - Option MACS (3^{ème} Année)

RAPPORT DE STAGE DE FIN D'ETUDE

FBH Associés

Analyse Quantitative :

Elaboration d'un modèle de risque de crédit

Encadrant en Entreprise:

M. Frédéric BOURY

Etudiant:

Djibril SARR

13/03/2018 – 13/09/2018



Table des matières

Table des figures.....	5
Table des Tableaux.....	6
Remerciements.....	7
Abstract.....	8
Introduction.....	9
I. Contexte du stage.....	10
1. Le cabinet FBH Associés.....	10
2. Contexte réglementaire : Les accords de Bâle.....	12
a. Présentation des accords de Bâle.....	12
b. Rappels sur Bâle II.....	12
c. Bâle III.....	14
d. Quelques mots sur l'ICAAP.....	15
3. La mission au sein de l'Institution Financière.....	16
II. Etat de l'art des modèles de Risque de Crédit.....	18
1. Définition du risque de crédit.....	18
2. Instruments de mesure du risque de crédit.....	19
a. Les Ratings.....	19
b. La distance au défaut.....	20
c. Les Matrices de Transitions.....	22
d. Exposition en cas de défaut (<i>Exposure At Default</i>).....	25
e. Perte en cas de défaut (<i>Loss Given Default</i>).....	25
f. Value at Risk (VaR).....	26
3. Les modèles de Risque de Crédit.....	28
a. Le modèle de Merton.....	28
b. Le modèle KMV.....	30
III. Le Modèle de Risque de Crédit.....	31
1. Conceptualisation du modèle.....	31
2. Le modèle de risque de crédit.....	34
3. La sélection des indicateurs macro-économiques du modèle.....	42
a. Choix du panel d'indicateurs macroéconomiques.....	42
b. Détermination du meilleur indicateur parmi le panel en mono-factoriel.....	45



c.	Détermination de la meilleure famille d'indicateurs en multifactoriel.....	57
4.	La gestion des chocs et Stress scénarios.....	64
IV.	Exemples de résultats.....	69
V.	Challenges et perspectives de perfectionnement du modèle.....	74
	Conclusion	75
	Annexes.....	76
A.	Régressions linéaires	76
B.	PCA.....	79
C.	K-means Clustering.....	81
D.	Recherche Documentaire sur le process de Notation de S&P	82
E.	Distribution NIG	86
	Références.....	87

Table des figures

Figure 1: Organigramme du Cabinet FBH Associés:.....	11
Figure 2: Mise en évidence de la Var et de la CVaR sur une distribution Gaussienne	27
Figure 3: Tirage des défauts dans la procédure 'Monte-Carlo'	33
Figure 4: Histogramme des R^2 : Inflation / Déflateur du PIB.....	47
Figure 5: Histogramme des R^2 : Accumulation nettes de passifs	47
Figure 6: Histogramme des R^2 : Epargne Intérieure brute	47
Figure 7: Histogramme des R^2 : Croissance du PIB / Hbt	47
Figure 8: Histogramme coefficients de corrélation: Dépenses nationales brutes	49
Figure 9: Histogramme coefficients de corrélation : Capitalisation boursière	49
Figure 10: Histogramme de la croissance du PIB/Hbt.....	50
Figure 11: Histogramme de la croissance du PIB/Hbt avec la NIG estimée	50
Figure 12: QQ-Plot contre la NIG estimée.....	50
Figure 13: Contribution des Indicateurs aux 4 dimensions principales:.....	53
Figure 14: Clustering sur les indicateurs macroéconomiques	53
Figure 15 : Evolution proportion R^2 optimaux fonction nombre d'indicateurs injectés:.....	56
Figure 16: Greedy Algorithm:	58
Figure 17: Ascending algorithm	60
Figure 18: Algorithme de sélection des indicateurs et de calcul des coefficients.....	61
Figure 19: Evolution des R^2 et R^2 ajustés en fonction du nombre d'indicateurs disponible:.....	63
Figure 20: Evolution de la VaR 1Y, 2Y et 3Y avec garanties pour tiers notées Ba1 (Moody's)	69
Figure 21: Evolution de la VaR 1Y, 2Y et 3Y avec garanties pour tiers notées Ba3(Moody's) .	69
Figure 22: Variation de la VaR après la considération des garanties.....	70
Figure 23: Effet d'un choc de LGD sur la VaR.....	71
Figure 24: Effet d'un choc positif sur la croissance du PIB	72
Figure 25: Evolution des ratings des Cps de l'IFI	73
Figure 26: Processus de l'algorithme k-means Clustering.....	81
Figure 27: Trame de rating des souverains de S&P	82

Table des Tableaux

Tableau 1: Correspondance entre ratings long termes internes, S&P, Moody's et Fitch.....	20
Tableau 2: Tableau de correspondance entre les ratings internes et la Distance au Défaut	21
Tableau 3: Matrice de Transition Etendue (23 notchs) Moody's par méthode de Cohorte	23
Tableau 4: Matrice de Transition Réduite (10 notchs) S&P par méthode de Duration.....	24
Tableau 5: Homogénéisation des indicateurs Tableau 1	44
Tableau 6: Homogénéisation des indicateurs Tab 2	45
Tableau 7: Performance des indicateurs en mono factoriel Tab 1.....	46
Tableau 8: Analyse de corrélations inter-pays à indicateurs constants.....	49
Tableau 9: Analyse de la qualité de fitting des distributions.....	51
Tableau 10: Approche particulières Clustering et PCA.....	54
Tableau 11: Meilleurs Indicateurs par Zone Géographique.....	55
Tableau 12: Exemple de trace de du greedy algorithm.....	57
Tableau 13: Occurrences des indicateurs dans les k-uplets résultants du greedy algorithm	59
Tableau 14: Evolution de la qualité des régression en fonction du nombre d'indicateurs disponible.....	63

Remerciements

Je remercie tout d'abord Monsieur Frédéric BOURY, associé fondateur du cabinet FBH-Associés. Sa bienveillance, ses conseils et sa disponibilité m'ont permis de mener à bien et en toute sérénité les travaux que j'ai le plaisir de présenter dans ce document.

Mes remerciements s'adressent ensuite au corps professoral de la spécialité Mathématiques Appliquées et Calcul Scientifique de l'école d'ingénieurs de Sup' Galilée. Plus particulièrement à Monsieur Olivier LAFITTE et Monsieur AUDUSSE, directeurs de la formation. Vos conseils et vos enseignements tout au long de notre formation nous ont guidés et assurément cela perdurera au-delà de l'école.

Enfin je remercie tous les proches qui m'ont soutenu pendant mon parcours. Particulièrement mes parents. Ma reconnaissance et ma gratitude à votre endroit sera éternelle.

Abstract

Dans un contexte réglementaire strict, où une mesure parfaite du capital économique est nécessaire, particulièrement en raison de l'ICAAP (Internal Capital Adequacy Assessment Process), le cabinet FBH-Associés a été chargé par une Institution Financière Internationale d'élaborer avec elle, un modèle interne de risque de crédit.

Le modèle adopté est un modèle multifactoriel. Il relie les variations de distance au défaut des contreparties aux variations d'un ou de plusieurs indicateurs macroéconomiques. Sa particularité est dans la corrélation qu'il impose entre les défauts des tiers non souverains et ceux des tiers souverains. Cette corrélation est dans ce modèle d'autant plus importante que la note du souverain est faible.

Le modèle utilise une famille de 5 indicateurs choisie par un couple d'algorithmes (ascendant et glouton).

En plus de donner en sortie, des paramètres de mesure du risque classique, EADs, nombres de défauts, VaR, CVaR, les travaux liés au modèle permettent aussi la simulation de stress scénarios incluant, notamment des chocs sur les LGDs, ou sur les indicateurs macroéconomiques du modèle.

Mots clés :

Analyse quantitative, Risque de crédit, VaR, CVaR, Distance au Défaut, Régressions Multifactorielles

Introduction

L'obtention du diplôme d'ingénieur en Mathématiques Appliquées et Calcul Scientifique (MACS) nécessite la validation d'un stage de fin d'étude d'une durée d'environ 6 mois. Ce stage est l'opportunité de mettre en pratique les théories et techniques étudiées lors des années de formation mais aussi de développer de nombreuses nouvelles compétences. Ce rapport de stage vient donc présenter les travaux qui ont été menés du 13 mars au 13 septembre lors de cette expérience professionnelle avec le cabinet de conseil FBH Associés.

Ces 6 mois de stage ont été consacrés à la construction et à l'implémentation d'un modèle interne de risque de crédit pour une Institution Financière Internationale (IFI). Les expositions au risque d'une IFI sont très différentes de celles des banques d'investissement privées. D'où la nécessité de construire un modèle propre à l'IFI, plutôt que d'utiliser les outils déjà disponibles. Ledit modèle est multifactoriel. Il met en relation les variations de distances aux défauts de tiers, aux variations de plusieurs indicateurs macro-économiques.

Il s'agira donc dans ce rapport de d'abord présenter le contexte de stage. Nous présenterons dans cette partie l'entreprise FBH Associés mais surtout nous introduirons la mission au sein de l'IFI, en expliquant pourquoi est-ce qu'il était nécessaire pour elle de construire un modèle interne. La deuxième partie de ce document sera celle qui s'intéressera au modèle à proprement parler. Un bref état de l'art des modèles de risque de crédit sera d'abord fait. Une attention particulière sera attribuée à la notion de distance au défaut qui sera d'abord analysée à travers le modèle de Merton avant qu'elle ne soit explicitée dans notre modèle interne. Celui-ci sera naturellement présenté en détail. Les équations qui le définissent seront explicitées. Aussi, les indicateurs macro-économiques qu'il utilise ainsi que leurs processus de sélection seront détaillés.

I. Contexte du stage

1. Le cabinet FBH Associés

Fort de maintenant, 27 années d'expériences, le cabinet a été fondé par Frédéric Boury et Denis Grandjean en 1992. Les deux associés fondateurs avaient déjà fait montre de leur expertise dans le conseil. L'activité du cabinet est concentrée en France mais ne s'y limite pas et des missions internationales, notamment en Afrique sont parfois assurées par les collaborateurs.

Le cabinet est à taille humaine (une quarantaine de consultants). Ce qui fait aujourd'hui sa force et qui est son caractère distinctif est qu'il place en son centre l'humain et non pas l'expansion du cabinet. C'est ainsi que voir les collaborateurs grandir et devenir des experts est plus valorisé chez FBH Associés que l'augmentation du nombre de collaborateurs.

Ainsi, l'évolution du cabinet s'est-elle faite de manière progressive et la diversification de ses activités est venue naturellement, en même temps que ses domaines d'expertise se sont développés. C'est ainsi que le cabinet a d'abord accompagné des clients en MOA (Maîtrise d'Ouvrage) et AMOA (Assistant à Maîtrise d'Ouvrage). Lorsque cette qualité d'accompagnement s'est bien développée, le cabinet a commencé à proposer son expertise en tant que cabinet de conseil.

Le cabinet travail avec des entreprises privées comme avec des instances publiques. Il intervient majoritairement dans les branches du secteur financier. Notamment :

- Les banques de financement et d'investissement ;
- Les banques de détail ;
- Les sociétés d'assurances ;

Néanmoins, elle intervient aussi dans d'autres secteurs. Parmi lesquels :

- Les industries ;
- Les sociétés de services.

Dans toutes ces entreprises et institutions publiques, les compétences nombreuses et variées des consultants du cabinet, justifient le grand nombre et la variété des missions menées par FBH Associées. Ainsi, les elles peuvent concerner en finance :

- Les risques, (réglementaires ou relatifs au crédit) ;
- La structuration ;
- Ou encore de l'ALM ("Assets and Liabilities Management"¹).

Un pôle digital du cabinet s'intéresse également à des questions telles que :

- L'ABC ("Activity Based Costing"²);
- La sociologie des organisations.

¹ Gestion actifs-passifs

² Méthode des coûts par activité

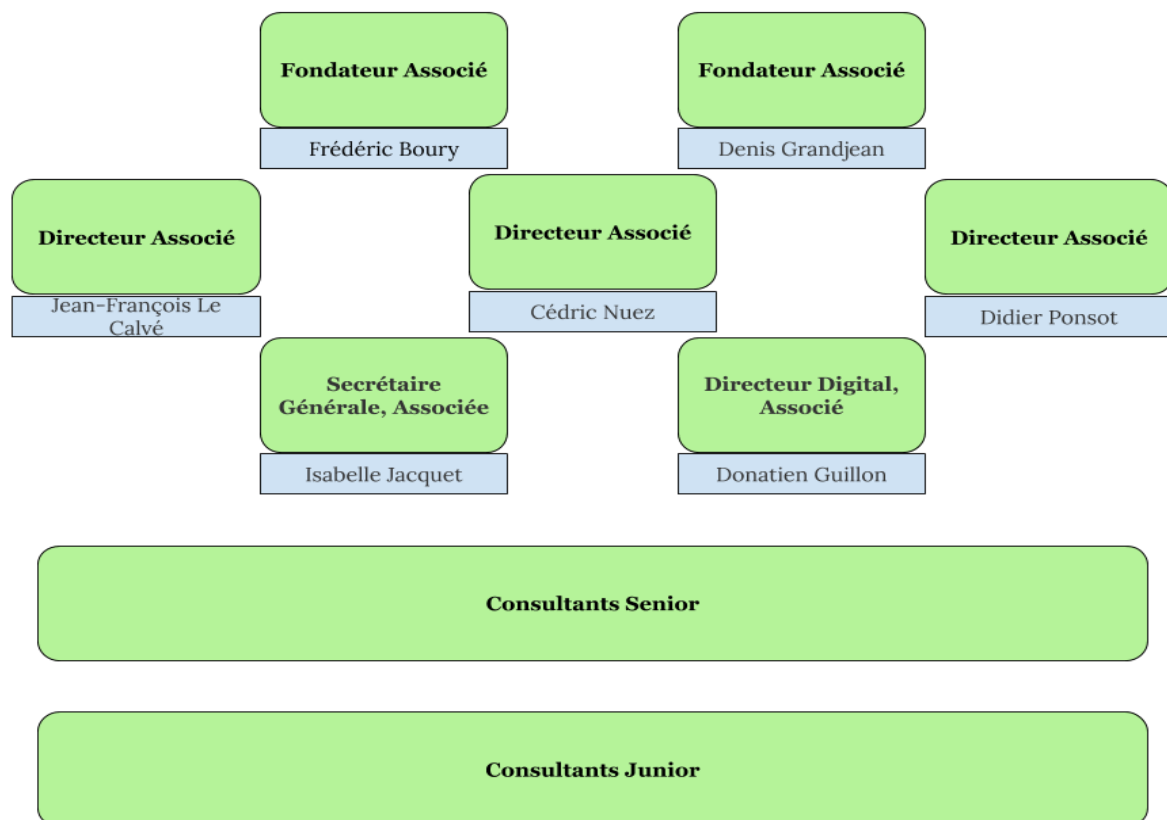


Figure 1: Organigramme du Cabinet FBH Associés:

2. Contexte réglementaire : Les accords de Bâleⁱⁱⁱ

La réglementation Bâloise accorde beaucoup d'importance au risque de crédit et encourage les banques à adopter un certain nombre de mesures pour se couvrir contre ce risque. C'est d'ailleurs dans ce contexte, que l'IFI élabore son modèle interne de mesure de capital économique. Ce dernier doit donc répondre à toutes les exigences réglementaires bâloises.

a. Présentation des accords de Bâle

Les accords de Bâle sont des textes de réglementation bancaire qui apporte des exigences minimales en fonds propres, dans le but de consolider la santé financière des établissements bancaires. Ces accords ont été établis par le comité de Bâle, un comité qui regroupe les principales banques centrales et qui traite régulièrement des problématiques de contrôle et de surveillance prudentiels.

Le premier accord de Bâle a été publié en 1988. Ayant pour objectif d'assurer la stabilité financière du secteur bancaire et d'installer les bases d'une réglementation plus stricte, cet accord a laissé paraître un certain nombre de lacunes au début des années 90. Ces dernières ont été comblées par la suite grâce à un second accord publié en 2004 : Bâle II.

b. Rappels sur Bâle II

La réforme Bâle II s'est articulée autour de trois piliers : l'exigence minimale en fonds propres, la surveillance prudentielle, ainsi que la discipline de marché.

Le premier pilier visait à imposer aux institutions bancaires un niveau minimal de fonds propres réglementaires relatifs aux risques opérationnels, risque de marché ainsi qu'au risque de crédit.

En ce qui concerne le risque de crédit, le nouvel accord de Bâle propose trois approches différentes pour calculer ces fonds propres :

- **Approche standard :**

Cette approche mesure le niveau minimum de fonds propres en appliquant des coefficients de pondération (*Risk-weight*) aux EAD des actifs du portefeuille.

Ces coefficients de pondération sont fixés par la réglementation et varient en fonction d'un certain nombre de paramètres notamment la catégorie d'actif ainsi que le *rating* externe (celui des agences de notations).

L'objectif de cette approche est de distinguer les actifs risqués des peu risqués. Les premiers seront d'avantage pondérés par le risque.

- **Approche IRB (*Internal Ratings Based*) simplifiée « *Foundation IRB*» :**

Contrairement à l'approche précédente, l'IRB simplifié utilise des *ratings* et des probabilités de défauts calculés en interne. Les autres paramètres restant fixés par les autorités de contrôle.

- **Approche IRB (*Internal Ratings Based*) avancée « *Advanced IRB* » :**

Cette approche utilise d'autres paramètres calculés en interne. En effet, si l'approche précédente ne s'appuyait que sur les *ratings* et PD internes, celle-ci s'appuie également sur les EADs et LGDs internes.

En somme, une des différences fondamentales entre ces trois approches réside dans **la source de données**. En effet, l'approche standard utilise essentiellement des données réglementaires et externes. L'approche IRB simplifiée utilise quant à elle des données internes et externes. Enfin, l'approche IRB avancée n'utilise que des données internes à l'exception des effets de diversification.

De même pour le risque opérationnel, trois méthodes sont proposées :

- **Approche BIA (*Basic Indicator Approach*)³:**

Cette approche mesure le niveau minimum de fonds propres en se basant sur le produit annuel brut (PAB). En effet, les fonds propres réglementaires au titre de l'année n sont calculés de la manière suivante : $FP_n = IE_n \times 15\%$ avec $IE_n = \sum_{i=1}^3 \max(PAB_{n-i}, 0)$.

- **Standardised Approach :**

Cette approche effectue le calcul de fonds propres réglementaires par ligne métier.

À chaque ligne métier est associé un coefficient de pondération β_j . Les fonds propres réglementaires sont calculés de la manière suivante : $FP_n = \sum_{i=1}^3 IE_{n,j} \times \beta_j$

Aussi, il est important de noter que les FP calculés par cette méthode sont généralement moins importants que les précédents.

- **Approche AMA (*Advanced Measurement Approach*)⁴:**

Cette approche repose sur l'élaboration d'un modèle interne qui doit être validé par les autorités de contrôle. La méthode avancée la plus utilisée à ce jour est la LDA « *Loss Distribution Approach*⁵ »

Le deuxième pilier sera évoqué dans [ce qui suit](#).

Le troisième pilier quant à lui avait pour objectif d'instaurer une communication financière transparente.

³ Approche élémentaire

⁴ Approches de mesures avancées

⁵ Approche de Distribution des pertes

c. Bâle III

Suite à la crise financière, de nombreuses insuffisances ont été constatées dans Bâle II, et les autorités de contrôle ont dû revoir en profondeur la réglementation bancaire.

La forme des accords demeure inchangée, Bâle III s'articule également autour des trois piliers cités précédemment, mais ces piliers ont été renforcés et leur contenu a été enrichi.

La crise des *Subprimes* a révélé que les exigences en fonds propres imposées dans Bâle II étaient insuffisantes. Pour pallier à cela, Bâle III exige, d'une part d'opter pour des fonds propres de meilleure qualité, en particulier la catégorie *Tier I*. Et d'autre part, d'augmenter le niveau minimal des fonds propres.

Il est important de noter qu'il existe trois catégories de fonds propres :

- **Tier I** : ce sont les fonds propres de base, ils sont généralement constitués d'actions, de titres subordonnés et de réserves.
- **Tier II** : cette catégorie désigne les fonds propres complémentaires, elle est généralement constituée de plus-values latentes, de provisions et de titres participatifs.
- **Tier III** : fonds propres sur-complémentaires destinés à couvrir les risques de marché.

Bâle III encourage les banques à augmenter la part des actions et des réserves dans leur fonds propres de base, et à installer des matelas de sécurité complémentaires dans lesquels elles pourront verser des fonds en cas d'expansion économique, et retirer des fonds en période de récession.

Aussi, la nouvelle réglementation est plus exigeante que la précédente, et impose aux banques de respecter les différentes limitations mises en place :

- le ratio de solvabilité (=endettement net/ FP) est passé de 8% (dans Bâle II) à 10,5% (dans Bâle III)
- le ratio de l'effet de Levier (= rapport entre le Tier I et les expositions totales) qui est désormais limité à 3%.

Un autre axe d'amélioration mis en évidence par la crise réside dans la gestion de la liquidité. Bâle III a créé de nouveaux ratios de liquidité (LCR « Liquidité à court terme » et NSFR « Liquidité à long terme ») afin d'aider les banques à vérifier qu'elles disposent bien de niveaux convenables et adéquats de liquidité à court (LCR) et long (NSFR) termes.

Enfin, Bâle III accorde beaucoup d'importance à la couverture du risque systémique et encourage les banques à adopter un certain nombre de mesures, notamment la prise en considération de l'effet de corrélation et de "contagion" entre les grandes institutions financières.

C'est d'ailleurs dans ce contexte, que l'IFI, élabore son modèle interne de capital économique. Ce dernier doit donc répondre à toutes ces exigences réglementaires et prendre en compte les effets précédemment cités.

d. Quelques mots sur l'ICAAP

Depuis Janvier 2016, la BCE a énoncé un certains nombres de règles concernant entre autres, la mesure du capital économique et l'optimisation de la gestion du portefeuille. Cela s'est fait à travers l'ICAAP (Internal Capital Assessment Adequacy Process, Bases II, Pilier 2). Ce dernier permet aux banques et institutions financières de s'assurer que leurs fonds propres sont bien en accord avec leurs profils de risque. Le capital calculé dans l'ICAAP est un capital économique (fonds propres économiques). Il est mesuré grâce à un modèle interne et se définit comme le montant nécessaire à la banque pour couvrir l'ensemble des risques qu'elle encourt. Depuis le deuxième accord de Bâle, il tente également de faire converger ces fonds propres économiques résultant des modèles internes, vers les fonds propres réglementaires imposés par les autorités de contrôle.

Le but est de pouvoir donner à la banque un outil de simulations des scénarios afin d'évaluer les impacts en risque et le besoin en capital qui en découle. Avant, le modèle qui est présenté, l'IFI suivait principalement une approche réglementaire qui ne pouvait pas être totalement satisfaisante pour capturer des phénomènes de concentration ainsi que des corrélations d'évènements. Par conséquent, en ligne avec la directive européenne CRD IV, l'IFI souhaite disposer de stratégies et de processus fiables, efficaces et exhaustifs pour évaluer et tracer en permanence le montant, le type et la répartition du capital approprié pour couvrir les risques auxquels l'établissement est ou pourrait être exposé.

3. La mission au sein de l'Institution Financière

La réglementation impose donc à l'IFI de disposer d'un outil interne lui permettant de mesurer le risque de crédit auquel elle est exposée. L'institution étant internationale et ayant considérée les particularités de son activité s'est rendue compte que les solutions progiciels classiques existant sur le marché, ne correspondraient pas à ses besoins.

Cela était en effet nécessaire en considérant que:

- Le nombre de ses contreparties est limité mais avec des expositions importantes réparties sur plusieurs zones géographiques en Europe ; La banque peut être exposée à des effets de concentration dans certains pays et secteurs ;
- Les contreparties de l'IFI, contrairement aux banques privées, sont principalement composées de souverains, d'entités publiques et d'institutions financières locales. En général celles-ci sont bien notées. Toutefois, elles sont exposées au risque systémique de leur pays et de leur zone géographique ;
- Les financements sont structurés avec plusieurs niveaux de garanties : cautions bancaires ou d'Etats, entités "mères",... Le risque sur ces opérations s'apprécie donc surtout en termes de "joint defaults"⁶ corrélés des contreparties et des garants.
- Les variations de notation des contreparties sont souvent atypiques. De plus, l'hypothèse gaussienne pour les transitions de notation et les corrélations aux facteurs d'environnement n'est pas vérifiée.

En outre, remarquons que l'IFI détient un important portefeuille de titres, le modèle doit tenir compte de la sensibilité de la valeur marchande des titres par rapport à la dégradation de la notation des émetteurs.

L'IFI a entamé à la fin de l'année 2017 une démarche sur la mesure de Capital Economique : Elle s'intéresse aux pires pertes possibles liées au risque de Crédit qu'elle puisse encourir à l'horizon d'une période de référence (1 à Y années) sur les encours d'actif et de hors-bilan caractérisés par leur profil d'écoulement (maturité) et le niveau d'exposition au risque de défaut de leur contrepartie/émetteurs.

Cette démarche est en priorité ciblée sur le risque de crédit et plus généralement sur les impacts de dégradation de la notation de ses contreparties.

La démarche s'inscrit dans le contexte suivant :

- Un suivi des contreparties avec un système de notation interne et avec les notations d'entreprises spécialisées (Moody's, Standard & Poors, Fitch Ratings) ([cf. partie II](#)) ;
- Une remontée des positions à savoir : Prêts amortissables, Engagements de financements signés, Portefeuilles titres, DAT (dépôts à terme) et Nostri ;
- Un Calcul des ratios réglementaires de risque Bâlois, fondés sur les notions d'exposition (EAD), de RWA (Risk Weighted Assets), de PD et de LGD.

⁶ Défaut en corrélation jointe



Lors du début du stage en mars 2019, une première version pilote du modèle existait déjà. Il a été ensuite amélioré et modifié de par sa construction mathématique même mais aussi par les données qui lui sont attribuées en entrées. Le modèle sera étudié en détail dans la [partie III](#).

II. Etat de l'art des modèles de Risque de Crédit

1. Définition du risque de crédit

Le risque de crédit est le risque qu'un emprunteur ne rembourse pas tout ou une partie de son crédit ou qu'il ne soit pas en mesure de respecter ses obligations contractuelles aux échéances prévues - par un contrat le liant au prêteur -. Cela entraîne en conséquence une interruption des flux de trésorerie et une augmentation des coûts de recouvrement. Si ce risque n'est pas adéquatement mesurer afin de couvrir les éventuels défauts, les montants prêtés et les intérêts dus viennent s'inscrire en perte.

L'évaluation du risque de crédit passe par une bonne connaissance du tiers. La banque s'appuie pour son évaluation du risque de crédit sur son expérience et éventuellement sur des outils statistiques (scores ou systèmes experts). Pour une opération donnée, le risque de crédit peut être minoré par la prise de garanties (cautions d'une personne ou d'une société de cautionnement, ...) mais également en incorporant dans le taux d'intérêt une marge dont la valeur dépend du niveau de risque.

L'évaluation du risque de crédit est faite en fonction de la capacité globale de l'emprunteur à rembourser un prêt selon ses conditions initiales. Bien qu'il soit impossible de savoir exactement qui manquera à ses obligations, une évaluation et une gestion adéquates du risque de crédit peuvent réduire la gravité des pertes. Afin de réduire au mieux les effets lors de l'occurrence de défauts de paiements, il est nécessaire d'évaluer et de gérer au mieux le risque de crédit grâce à des modèles pertinents.

2. Instruments de mesure du risque de crédit

a. Les Ratings⁷

Un rating est une évaluation quantifiée de la solvabilité d'un emprunteur. Celui-ci peut être un particulier, une société, une autorité étatique ou provinciale ou encore un gouvernement souverain. Cette mesure de la solvabilité peut être faite sur une base générale ou à l'égard d'une dette ou d'une obligation financière particulière. Elle cherche à prédire la vraisemblance qu'une dette soit remboursée. Implicitement donc elle donne une information sur la probabilité qu'une dette soit remboursée. Les notations peuvent concerner le court, le moyen ou le long terme.

Les ratings peuvent être attribués par chaque entité financière (publique ou privée) à travers un système de notation interne. Mais aussi, des entreprises ont fait des ratings leur spécialité. Parmi elles, les plus connues sont certainement, Moody's Investors Service, Fitch Ratings, Standard & Poor's Financial Services LLC⁸.

Moody's	S&P	FITCH	Interne IFI	Descriptif	Type
Aaa	AAA	AAA	10	Première Qualité	Prime
Aa1	AA+	AA+	9,5	Haute Qualité	
Aa2	AA	AA	9		
Aa3	AA-	AA-	8,5		
A1	A+	A+	8	Qualité Moyenne Supérieure	
A2	A	A	7,5		
A3	A-	A-	7		
Baa1	BBB+	BBB+	6,5	Qualité Moyenne inférieure	
Baa2	BBB	BBB	6		
Baa3	BBB-	BBB-	5,5		
Ba1	BB+	BB+	5	Spéculatif	Non Prime
Ba2	BB	BB	4,5		
Ba3	BB-	BB-	4		
B1	B+	B+	3,5	Très Spéculatif	
B2	B	B	3		

⁷ Notations

⁸ Limited liability company (Compagnie à Responsabilité Limitée)

B3	B-	B-	2,5	
Caa1	CCC+	CCC	2	Risque élevé
Caa2	CCC	CC	1,5	Ultra spéculatif
Caa3	CCC-	C	1	Défaut avec faible possibilité de recouvrement
Ca	R	DDD	0,5	
C	SD	DD	0	Défaut sélectif ⁹
D	D			Défaut

Tableau 1: Correspondance ¹⁰entre ratings long termes internes, S&P, Moody's et Fitch

Notons aussi l'existence d'un dernier état, celui des tiers non notés, N.R ("Not Rated").

b. La distance au défaut

Au même titre que les ratings, il existe différents outils servant à évaluer la solvabilité d'un tiers. La mesure de la distance au défaut (DD), parmi les mesures fondées sur le marché, a gagné en importance au cours des dernières années. Cela s'explique partiellement par son industrialisation à travers le modèle KMV de Moody's (cf [partie III.3](#)).

Le sens de la distance au défaut est assez intuitif puisqu'elle tente en effet, de mesurer l'écart qu'il y a entre le niveau de solvabilité d'une contrepartie et le défaut. Elle intervient également dans le modèle de Merton (cf [partie III. 4](#)) qui modélise les capitaux propres d'une entreprise comme une option d'achat sur la valeur de ses actifs.

La mesure de la distance au défaut est donc associée à la probabilité que le marché de l'actif d'une entreprise soit inférieur à la valeur de sa dette. Cette définition se généralise naturellement aux municipalités, souverains et autres entités publiques.

Au sein de l'IFI, une transformation linéaire permet de passer des notes internes aux distances au défaut.

Dans les équations propres au modèle interne à venir, la DD, sera notée Z

⁹ Selon S&P, le défaut sélectif (SD) intervient lorsque, l'emprunteur n'a pas honoré une certaine partie de ses obligations ou une émission spécifique mais qu'il continue de payer ses autres types d'emprunts dans les temps

¹⁰ Source : Magyar Nemzeti Bank (Banque Centrale Hongroise)



Interne IFI	DD
10	7
9,5	6,5
9	6
8,5	5,67
8	5,33
7,5	5
7	4,67
6,5	4,33
6	4
5,5	3, 67
5	3,33
4,5	3
4	2,67
3,5	2,33
3	2
2,5	1,5
2	1
1,5	0,8
1	0,6
0,5	0,4
0	0,2
	0

Tableau 2: Tableau de correspondance entre les ratings internes et la Distance au Défaut

c. Les Matrices de Transitions

Les matrices de transitions ("credit migration matrices"), définissent les probabilités pour un tiers de passer d'un rating R_1 vers un rating R_2 différent ou similaire au rating R_1 . Ainsi, ce sont des matrices P carrés de dimensions le nombre de "notches"¹¹ de ratings. On définit ainsi :

$$P_{ij} = P(X_2 = R_j | X_1 = R_i)$$

La probabilité de variation de ces ratings peut être définie en fonction de l'horizon de couverture souhaité (annuelle, trimestrielle, ...).

Remarquons que les probabilités de transitions peuvent différer d'un secteur (banques, souverains, municipalités, ...) à un autre ; mais aussi d'une zone géographique à une autre. Il convient d'affiner autant que nécessaire les critères de regroupement de tiers afin de calculer les matrices de transition propres à un groupe bien précis.

Principalement deux méthodes de calcul des matrices de transitions existent.

Méthode des cohortes (ou des générations):

Cette méthode est la plus classique et la plus intuitive. Elle procède tout simplement en un décompte de la proportion de contreparties passées d'un rating i vers un rating j

$$P_{ij} = \frac{\sum_{t \in [1, T]} N_{ij}(t)}{\sum_{t \in [1, T]} N_i(t)}$$

- N_{ij} est le nombre de contreparties qui passent de la note i à la note j entre les deux périodes choisies Y et $Y + 1$ (par exemple, année Y et année $Y + 1$) ;
- N_i est le nombre de contreparties notées i à l'année Y .
- T est la profondeur d'historique de notation disponible (les notations peuvent exister par exemple entre les années 1970 et 2000, on a dans ce cas, $T = 30$, lorsque $t = 2$, on considérera toutes les variations de notations entre 1971 et 1972).

On constatera que cette approximation correspond à l'estimateur de maximum de vraisemblance lorsqu'on suppose que la variation des ratings suit une Chaîne de Markov (CDM) (ce que l'on sait ne pas être vrai).

Cette méthode qui pourtant est très souvent employée présente trois inconvénients principaux.

- Le premier est que cette technique ne prend pas en compte les transitions intermédiaires. Ainsi, si par exemple la matrice de transition 1 année est voulue, aucune des transitions entre deux mois - ou entre deux jours en fonction de la fréquence des ratings - n'est considérée. Des informations sont donc négligées et inexploitées.
- Ensuite, un autre désavantage est que pour cette méthode, toutes les contreparties qui sont passés en N.R. au terme de la période, ne sont pas incluses

¹¹ Niveaux

dans le calcul des probabilités de transition. Les conséquences deviennent donc les mêmes que pour le point précédent.

- Enfin, le dernier inconvénient est que, puisque cette méthode est basée sur les occurrences, elle n'est pas capable de fournir des probabilités de transitions strictement positives pour des événements non observés. Ainsi, si par exemple, aucun défaut n'a été constaté dans l'échantillon considéré, les probabilités de défaut sont systématiquement nulles.

	Aaa	Aa1	Aa2	Aa3	A1	A2	A3	Baa1	Baa2	Baa3	Ba1	Ba2	Ba3	B1	B2	B3	Caa1	Caa2	Caa3	Ca	C	D
Aaa	0.98	0.01	0.01	0.00	0.00	0.00	0.00	0.00	-	-	0.00	-	-	-	-	-	-	-	-	-	-	-
Aa1	0.03	0.89	0.03	0.03	0.00	0.01	0.00	-	-	0.00	-	-	-	-	-	-	-	-	-	-	-	-
Aa2	0.02	0.04	0.89	0.03	0.01	0.01	0.00	0.00	-	0.00	-	-	0.00	-	-	-	-	-	-	-	-	-
Aa3	-	0.01	0.04	0.89	0.03	0.03	0.00	-	0.00	-	-	-	-	-	-	-	-	-	-	-	-	-
A1	0.00	0.01	0.01	0.04	0.83	0.09	0.01	0.01	0.00	0.00	-	-	0.00	0.00	-	-	-	-	-	-	-	-
A2	0.01	0.03	0.03	0.07	0.05	0.79	0.02	0.01	0.00	0.00	-	-	-	-	-	-	-	-	-	-	-	-
A3	0.01	0.01	0.00	0.03	0.02	0.05	0.84	0.02	0.02	0.00	0.00	-	-	-	-	-	-	-	-	-	-	-
Baa1	0.01	-	0.02	0.01	0.03	0.04	0.02	0.78	0.06	0.03	0.01	-	0.00	-	-	-	-	-	-	-	-	-
Baa2	-	0.00	0.02	0.04	0.01	0.09	0.02	0.01	0.75	0.04	0.01	0.00	0.01	-	-	-	-	-	-	-	-	-
Baa3	0.00	0.01	0.04	-	0.01	0.00	0.02	0.01	0.03	0.84	0.03	0.01	0.01	-	-	-	-	-	-	-	-	-
Ba1	0.01	-	-	-	0.03	0.02	0.01	0.01	0.01	0.05	0.80	0.04	0.01	0.01	-	-	-	-	-	0.01	-	-
Ba2	-	-	0.03	-	0.01	-	0.01	0.01	0.03	0.04	0.09	0.61	0.11	0.03	0.01	0.01	0.01	-	-	-	-	-
Ba3	-	0.01	0.02	0.01	0.08	0.04	0.01	-	0.01	0.01	0.01	0.04	0.68	0.07	0.02	-	-	-	-	-	-	-
B1	-	-	-	-	0.03	-	0.03	-	-	-	0.03	0.01	0.06	0.81	0.03	-	-	-	0.01	-	-	-
B2	-	-	-	-	-	0.05	-	-	-	-	0.12	0.02	-	-	0.69	0.10	0.02	-	-	-	-	-
B3	-	-	-	0.02	0.02	-	-	-	-	-	0.02	0.04	0.02	0.06	0.75	0.04	0.02	-	-	-	-	-
Caa1	-	-	-	-	-	-	-	-	0.03	-	-	0.03	0.07	-	0.07	0.03	0.69	-	0.03	0.03	-	-
Caa2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.50	0.50	-	-	-
Caa3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.20	0.20	-	0.40	-	0.20	-
Ca	-	-	-	-	-	-	-	0.25	-	-	-	-	-	-	-	0.25	-	-	-	0.50	-	-
C	-	-	-	-	-	-	-	-	-	-	1.00	-	-	-	-	-	-	-	-	-	-	-
D	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.00

Tableau 3: Matrice de Transition Etendue (23 notches) Moody's par méthode de Cohorte

Méthode de Duration :

Cette méthode permet de contourner chacun des inconvénients de la méthode de générations. Elle consiste en le calcul d'une matrice génératrice Γ puis en l'application d'une exponentielle de matrice à celle-ci

$$P = e^{\Gamma}$$

On rappelle qu'une matrice génératrice est une matrice vérifiant :

$$\forall i \neq j, M(i, j) \geq 0$$

$$\forall i, \sum_j M(i, j) = 0$$

La matrice génératrice est la matrice Γ vérifiant :

$$\begin{cases} \gamma_{ij} = \frac{n_{ij}(T)}{\int_0^T Y_i(s) ds}, i \neq j \\ \gamma_{ii} = -\sum_{i \neq j} \gamma_{ij} \end{cases}$$

- $Y_i(s)$ est le nombre de tiers notés i à l'instant s .
- $n_{ij}(T)$ est le, nombre total de transitions de i vers j
- $\int_0^T Y_i(s) ds$ qui dans les calculs sera approximé par une somme permet de considérer toutes les transitions intermédiaires intervenues avant la fin de la période ciblée pour la matrice de transition.

	AAA	AA	A	BBB	BB	B	CCC	CC	C	D
AAA	0.861	0.098	0.006	-	-	-	0.002	-	-	-
AA	0.002	0.855	0.105	0.005	-	-	-	-	-	-
A	0.000	0.019	0.875	0.059	0.002	0.000	-	-	-	0.000
BBB	-	0.001	0.046	0.858	0.034	0.004	0.001	0.000	-	0.001
BB	-	-	-	0.054	0.753	0.076	0.004	0.001	-	0.003
B	-	-	0.001	0.003	0.067	0.737	0.040	0.004	-	0.023
CCC	-	-	-	0.005	-	0.187	0.444	0.027	-	0.198
CC	-	-	-	-	-	0.044	0.087	0.044	-	0.783
C	-	-	-	-	-	-	-	0.200	0.970	0.100
D	-	-	-	-	-	-	-	-	-	1.000

Tableau 4: Matrice de Transition Réduite (10 notchs) S&P par méthode de Duration

Comparaison de deux Matrices :

Les matrices de transitions permettant d'évaluer le niveau de risque au sein d'un périmètre de tiers, il est nécessaire de savoir comparer d'une période à une autre (typiquement d'une année à une autre) deux matrices afin de mesurer l'évolution du risque inhérent au périmètre en question.

Ou même, il peut être nécessaire d'évaluer le risque d'un périmètre à un autre.

Une des méthodes consiste en l'utilisation de la décomposition en valeur singulière (SVD) de la 'matrice de mobilité'. En remarquant que la matrice identité caractériserait un cas fictif où il n'y aurait aucune transition.

Ainsi, on définit une matrice de mobilité $\tilde{P} = P - Id$

La métrique de comparaison M_{SVD} est finalement la moyenne des valeurs singulières de la matrice de mobilité.

$$M_{SVD}(P) = \frac{1}{N} \sum_i^N \sqrt{\lambda_i(\tilde{P}^t \tilde{P})}$$

Plus le coefficient $M_{SVD}(P)$ est élevé, plus il y a de variabilité au sein de ce périmètre. Malheureusement, cette technique ne permet pas de comparer les tendances des périmètres (est-ce que les notes s'apprécient ou se déprécient) mais seulement le niveau de variabilité

d. Exposition en cas de défaut (*Exposure At Default*)

L'exposition en cas de défaut (EAD) correspond au montant non réglé par l'emprunteur au moment du défaut. C'est le capital restant dû et intérêts à recevoir au moment du défaut.

e. Perte en cas de défaut (*Loss Given Default*)

Lorsqu'une contrepartie est en défaut, son actif sert en premier temps au recouvrement de ses dettes. Si, l'actif de la contrepartie en défaut ne suffit pas pour rembourser toutes ses dettes, une partie de la dette sera remboursée (taux de recouvrement, TDR), et l'autre sera perdue (LGD). Ainsi, la LGD correspond au pourcentage de la dette qui n'a pas pu être remboursée. On a naturellement $LGD = 1 - TDR$.

f. Value at Risk (VaR)

La VaR est un outil de gestion des risque qui permet d'indiquer des niveaux de pertes sur un certain horizon temporel et à un seuil de confiance fixé¹². La VaR n'est autre qu'un quantile de la distribution de perte à un certain seuil de confiance. En effet, si on considère un seuil de confiance à 99%, la VaR à un an correspond au quantile $q_{99\%}$ de la distribution des pertes (et profits) à horizon un an. Cette valeur correspond au **niveau de perte qui ne devrait pas être dépassé dans 99% des cas**.

Le calcul de la *Value at risk* peut se faire selon trois approches différentes :

L'approche historique :

Cette approche consiste à constituer la distribution des pertes à partir des observations historiques. Une fois la distribution obtenue, nous déterminons la VaR en prenant le quantile de cette distribution à un certain seuil.

L'inconvénient de cette approche réside dans le fait que l'on utilise les observations historiques qui ne sont pas forcément représentatives de la distribution des pertes réelles

L'approche paramétrique :

Cette approche consiste à utiliser une loi de distribution paramétrique des pertes (et profits). Pour déterminer cette loi, on utilise un certain nombre de facteurs de risque (comme le rendement des actifs) et on choisit la loi qui correspond le mieux aux qualités distributionnelles de ces facteurs.

Le principal inconvénient de cette approche est qu'elle suppose la connaissance des distributions des facteurs de risque, ce qui n'est pas souvent le cas.

L'approche par Monte-Carlo :

C'est cette méthode qui sera utilisée dans ce rapport.

Cette approche consiste à effectuer un grand nombre de simulations dans le but d'obtenir la distribution des pertes/gains liés à notre portefeuille d'étude.

Ces simulations prennent en considération l'évolution des facteurs de risque et variables de marché ayant un impact sur l'exposition au risque.

Malgré le fait qu'elle soit considérée comme la plus « sophistiquée » des trois méthodes, elle présente un inconvénient majeur. En effet, La fiabilité de la distribution de pertes obtenue

¹² Ces niveaux de pertes ne devraient être atteints qu'avec une probabilité donnée (1 - seuil de confiance)

dépend essentiellement de la fiabilité des modèles utilisés pour faire évoluer les facteurs de risque et la fiabilité de leur calibration (généralement historique...) qui peut avoir pour effet de sous-estimer les facteurs de risque et implicitement les pertes du portefeuille d'étude.

Aussi, l'une des limites de la *Value at Risk* est **qu'elle néglige les pertes subies au-delà de sa valeur**. En effet, en déterminant la VaR d'une distribution de perte, nous quantifions la pire perte possible à un certain niveau de confiance, mais nous n'avons pas d'informations concernant l'amplitude des pertes subies au-delà de cette VaR.

Pour pallier à cela, nous pouvons calculer une Expected Shortfall ¹³ou *Conditional Value at Risk* ¹⁴(CVaR). Celle-ci est l'espérance de la perte lorsque cette dernière dépasse la VaR. On se focalise donc sur la queue de la distribution (figure 2) d'où son autre appellation Tail-VaR.

$$CVaR = E[-X | X < -VaR_\alpha(X)]$$

$$CVaR = \frac{1}{1-\alpha} \int_{-\infty}^{-VaR} xp(x)dx$$

$p(x)dx$ étant la densité de probabilité d'obtenir un retour de x

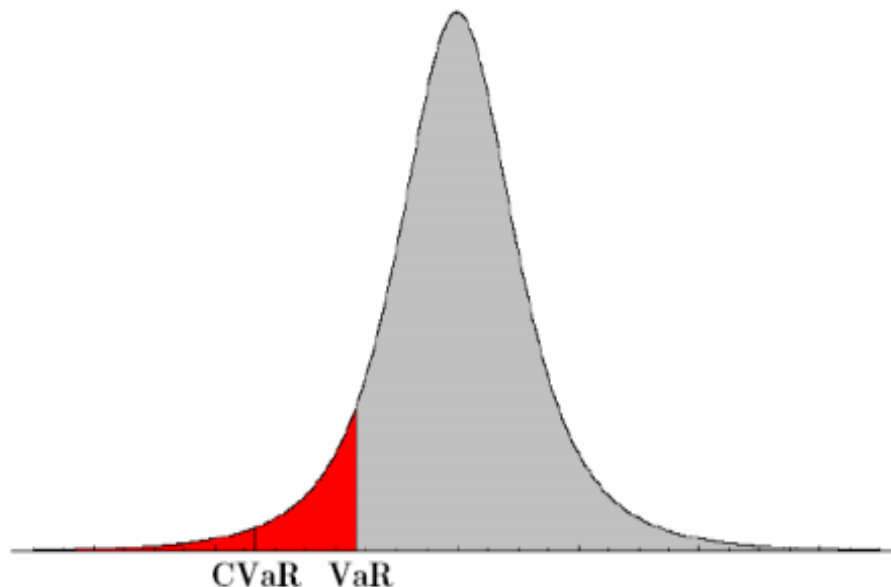


Figure 2: Mise en évidence de la VaR et de la CVaR sur une distribution Gaussienne

¹³ Déficit Attendu

¹⁴ Valeur en jeu Conditionnelle

3. Les modèles de Risque de Crédit

a. Le modèle de Mertonⁱⁱⁱ

Le modèle de Merton part d'une structure assez simplifiée de la dette. Il admet que la valeur totale A_t des actifs de la firme suit un processus géométrique

$$dA_t = rA_t dt + \sigma A_t dW_t, \quad A_0 > 0$$

- r est la moyenne de rappel
- σ est la volatilité

Le modèle fait quelques assumptions de plus :

- Le prix de liquidation de la compagnie correspond à sa valeur. En d'autres termes, il n'y a pas de charges de faillites
- La dette et les capitaux propres ("equity") sont des actifs dans un marché sans coûts de transaction ("frictionless market")
- La dette est constituée d'une seule obligation ayant une valeur nominale K et une maturité T .

À l'échéance, si la valeur totale des actifs est supérieure à la dette, celle-ci est intégralement payée et le solde est réparti entre les actionnaires. Toutefois, si $A_T < K$, un défaut advient: les obligataires exercent leur droit de liquider l'entreprise et de recevoir la valeur de liquidation (égale à la valeur totale de l'entreprise puisqu'il n'y a pas de frais de faillite) en lieu et place de la dette. Les actionnaires ne reçoivent rien dans ce cas, mais en vertu du principe de la responsabilité limitée ne sont pas tenus d'injecter des fonds supplémentaires pour payer la dette.

De ces simples observations, nous voyons que les actionnaires ont un flux de trésorerie à T égal à $(A_T - K)^+$. Ainsi, les fonds propres peuvent donc être considérés comme une option d'achat européenne sur les actifs de l'entreprise. D'autre part, le détenteur de l'obligation reçoit $\min(A_T, K)$.

Nous pouvons également écrire que la probabilité de défaut s'écrit :

$$P_t[\tau = T] = P_t[A_T \leq K]$$

La valeur théorique de l'entreprise est donc en appliquant les résultats classiques pour les calls européens est sous la probabilité risque neutre Q (où $e^{-rt}(A_t)_t$ est une martingale) :

$$\begin{aligned} E_t &= E^Q[e^{-r(T-t)} (A_T - K)^+] \\ &= A_t N(d_1) - e^{-r(T-t)} K N(d_2) \end{aligned}$$

Avec :

$$\begin{cases} d_1 = \frac{\log\left(\frac{A_t}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}} \\ d_2 = \frac{\log\left(\frac{A_t}{K}\right) + \left(r - \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}} \end{cases}$$

Remarquons quelques limites au modèle de Merton ^{iv}:

- La structure non stationnaire de la dette conduit à la cessation des opérations à une date fixe. Le défaut ne peut survenir qu'à cette date. Notons que le modèle de Merton peut néanmoins être étendu au cas des obligations de différentes échéances.
- La valeur de l'entreprise et ses paramètres ne sont naturellement pas directement observables.
- Les taux d'intérêt devraient être considérés comme stochastiques.

Dans le modèle de Merton, la probabilité de défaut correspond à $N(-d_2)$. Et celle-ci est l'application de la fonction de répartition de la loi Normale centrée réduite à l'opposé de la distance au défaut déjà évoquée.

Ainsi, dans le modèle de Merton, nous avons :

$$DD = \frac{\log\left(\frac{A_t}{K}\right) + \left(r - \frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}}$$

b. Le modèle KMV^{15v}

À l'instar du modèle de Merton, le modèle KMV permet également de calculer la probabilité de défaut en fonction des fondamentaux d'une entreprise (dettes, valeurs, ...). L'approche KMV utilise d'ailleurs en partie le modèle de Merton. Dans la dernière étape, KMV utilise la distance au défaut, puis se concentre sur l'expected default frequency ¹⁶(EDF) (qui en réalité correspond à la probabilité de défaut) en fonction des tendances historiques.

Même si le processus géométrique utilisé par KMV est le même que pour Merton, des différences sont à souligner.

- **Point de défaut** : Une autre différence importante est que le "point de défaut" du modèle KMV n'est pas le même que le point où la valeur des actifs tombe en dessous de la valeur de la dette totale, contrairement au modèle de Merton. En effet, le défaut se produit lorsque la valeur des actifs de l'entreprise se situe entre la valeur de la dette à court terme et celle de la dette totale. Il est en conséquence possible de ne pas avoir de défaut de paiement même si la valeur des actifs est tombée à un niveau inférieur à la dette totale.
- **Valeur de l'entreprise** : Plutôt que d'essayer d'estimer la valeur de l'entreprise directement à partir de données de bilan détaillées (une procédure très sensible aux hypothèses et à la méthode), KMV déduit A_t de la valeur de la dette (qui est tirée du bilan) et des capitaux propres.
- **Etalonnage** : La façon appropriée d'estimer les paramètres du modèle et le processus de détermination de la valeur de l'entreprise, après l'observation d'une série temporelle de capitalisations boursières $\{\hat{E}_{t_1}, \hat{E}_{t_2}, \dots\}$. utilise le théorème de Duan :

$$(\hat{\sigma}, \hat{\mu}) = \operatorname{argmax}_{\sigma, \mu} L_E(\hat{E}_{t_1}, \hat{E}_{t_2}, \dots, \hat{E}_{t_N}; \sigma, \mu)$$

$$\text{et } \hat{A}_{t_i} = F_i^{-1}(\hat{E}_{t_i}, \hat{\sigma})$$

$$\text{avec } F_i = E^Q[e^{-r(t-t_i)}(A_t - K)^+]$$

De manière assez semblable au modèle de Merton, on introduit une distance au défaut

$$DD_t = \frac{\log\left(\frac{A_t}{K}\right)}{\sigma}$$

Et finalement $EDF_t = N(DD_t)$

¹⁵ Le modèle porte le nom de ceux qui l'ont établi : Kealhofer, McQuown et Vasicek

¹⁶ Fréquence de défauts attendue

III. Le Modèle de Risque de Crédit

1. Conceptualisation du modèle

Le modèle cherche à mesurer les impacts de la dégradation de la notation de ses contreparties sur son capital économique.

Toute la particularité du modèle réside en le fait que les problématiques et les expositions aux risques sont [très différentes pour cette IFI](#). En effet, ses contreparties sont constituées de bien plus d'entités publiques que privées. De plus, très souvent les garanties des contreparties de la Banque sont des Etats. Il convient donc de pouvoir relier adéquatement l'évolution des variations de DDs des tiers non souverains à celles des Etats.

Dans sa conceptualisation, notre modèle de risque de crédit cherche à mettre en relation de manière linéaire, les variations de distance au défaut des contreparties de la banque aux variations d'indicateurs macro-économiques. La notion de distance au défaut ainsi que nous l'avons déjà vu dans l'état de l'art est en effet depuis le modèle de Merton, reliée à un facteur systémique, en l'occurrence le niveau d'endettement. L'idée du modèle est donc de relier de manière affine plusieurs facteurs systémiques, aux variations de distance au défaut. En l'occurrence, les facteurs systémiques correspondront à des indicateurs macroéconomiques.

Néanmoins, en pratique, nous ne nous intéressons pas à établir un modèle économique explicatif de la DD. En effet ainsi qu'il a été énoncé [dans ce qui précède](#), nous retiendrons uniquement qu'il existe une transformation affine entre le rating d'une contrepartie et sa distance au défaut Z_i d'une part et que cette variable Z_i est supposée vérifier ces assertions

$$\begin{aligned}\Delta Z_i &= \mu_{zi} \Delta t + \sigma_{zi} \Delta W \\ \Delta W &\sim N(0, 1)\end{aligned}$$

1

Dans cette expression,

- ΔZ désigne la variation de distance au défaut ;
- μ_{Z_i} et σ_{Z_i} désignent respectivement les paramètres de drifts et de volatilité propre aux variations de ratings de la contrepartie (i);
- Δt est l'intervalle de temps considéré ;
- ΔW est un aléa gaussien centré réduit.

Il s'agira donc de faire des simulations pluriannuelles en Monte-Carlo de l'évolution des notations (de la distance au défaut en fait) et donc a fortiori de simuler les passages en défaut des contreparties et émetteurs. La simulation est étendue également aux tiers intervenant comme garants pour le compte des bénéficiaires, mais aussi aux sociétés mères. Ces deux derniers sont activés en cas de défaut intervenu sur les contreparties.

L'évolution des notations est également contrôlée par l'application de matrices de transition annuelles en fonction du secteur d'activité des contreparties. En raison du caractère non gaussien des distributions, les variations de distance au défaut sont toujours obtenues au final par l'utilisation des matrices de transition : il n'y a pas de calcul opéré qui donne directement la variation de distance au défaut par simple propagation gaussienne. Il est nécessaire de se repositionner sur les distributions observées via les matrices de transitions.

Notons également que le modèle et notamment la partie s'intéressant aux [chocs factoriels](#) intègre également les corrélations entre les variations des indicateurs macroéconomiques entre les pays à travers un modèle de copule gaussienne.

Enfin, le modèle interne de l'IFI a la particularité de prendre en compte une corrélation de défaut non seulement par rapport au facteur systémique mais aussi par rapport à la situation de l'Etat du pays.

Résumé de la conceptualisation du modèle

Le modèle est basé sur une simulation pluriannuelle Monte-Carlo de la variation de notation de tiers et, a fortiori, des défauts de paiement :

- Nous générons l'évolution pluriannuelle de la notation des tiers : contreparties et émetteurs, mais aussi garants ou sociétés holding ;
- En cas de premier défaut de la contrepartie, les garanties sont activées et nous cherchons à voir l'occurrence de "défauts persistantes" (si les garanties sont en défaut aussi). Dans une situation de joint default, nous appliquons la LGD à l'EAD de toutes les contreparties impliquées.

Le modèle procède à une diffusion stochastique de la notation de chaque tiers en fonction de ses paramètres (drift, volatilité) et de sa corrélation avec les facteurs macro-économiques :

- Les variations ne suivent pas un processus gaussien ordinaire, mais sont basées sur la distribution réelle provenant des matrices de transitions ;
- Les variations de notations de tiers ont une composante factorielle relative à sa corrélation avec l'indicateur macroéconomique pays/secteur et une composante idiosyncratique.
- Les composantes factorielles des différents pays/secteurs sont liées par un modèle de copule.
- Cependant, au-delà d'une approche de modèle factoriel ordinaire, le modèle de l'IFI intègre également les effets systémiques, traités comme des "comportements corrélés" de tiers.

La composition des deux processus (celui du facteur et celui lié à la composante idiosyncratique) conduit à donner à la variation de distance au défaut une expression (de l'équation 1) comportant trois dimensions: une dimension temporelle et deux composants stochastiques indépendants.

Autrement dit, l'équation (1) se réécrit : $\Delta Z_t = \mu_{zi} \Delta t + \sigma_{zi} [a. \Delta W_1 + b. \Delta W_2]$

En résumé, le modèle de l'IFI comporte un add-on pour modéliser la corrélation spécifique entre les tiers, en plus des corrélations à facteur unique : certains des tiers sont appelés des "master third parties"¹⁷ et leur comportement influence un ensemble appelé "ordinary third parties"¹⁸.

Nous avons utilisé cette approche pour créer une corrélation supplémentaire entre l'État et tous les tiers d'un pays, afin d'aborder le risque systémique.

- L'idée principale est que la corrélation entre les tiers ordinaires et l'État augmente et peut potentiellement l'emporter sur tous les autres effets en cas de crise de solvabilité (se traduisant par une dégradation drastique de la solvabilité de l'État) ;
- Par cette approche, tous les tiers du pays (contreparties et garants) adoptent un comportement lié qui devient prédominant pour déclencher des défauts communs corrélés.

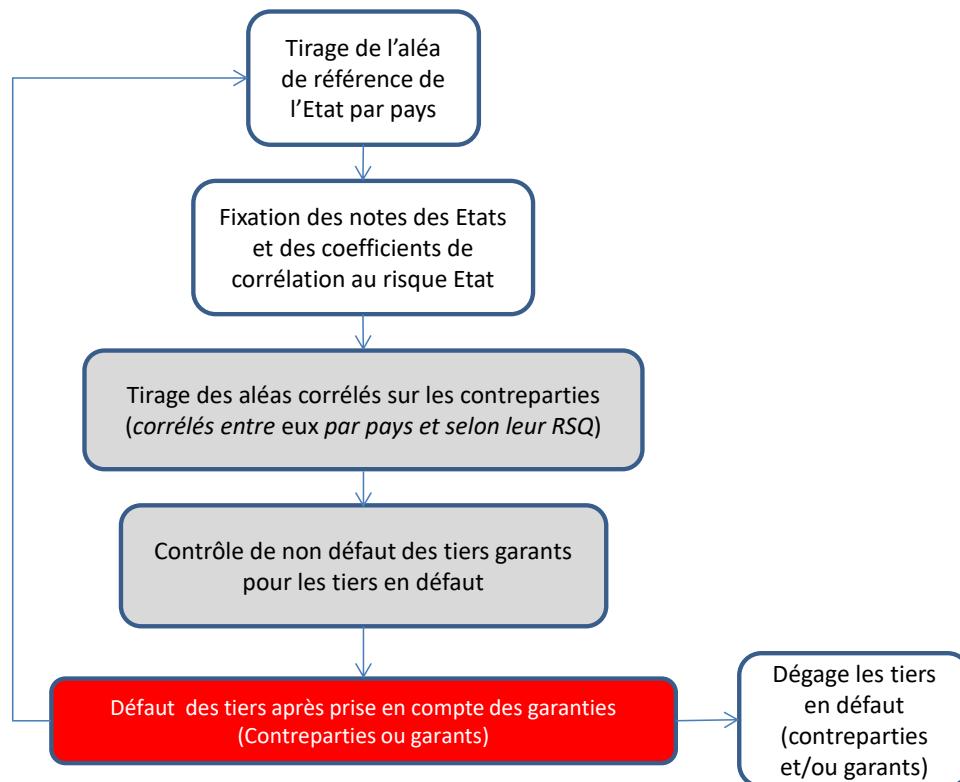


Figure 3: Tirage des défauts dans la procédure 'Monte-Carlo'

¹⁷ Tiers dominants

¹⁸ Tiers ordinaires

2. Le modèle de risque de crédit

Dans ce qui suit, nous utilisons des résultats sur les régressions linéaires qui seront développés en [annexe](#).

Nous admettons donc qu'il existe à un bruit Gaussien près une transformation affine permettant de passer des variations de ratings aux variations d'indicateurs macroéconomiques.

Remarquons que le modèle a été complexifié en deux étapes. Les travaux qui seront présentés dans ce qui suit concernent la dernière évolution du modèle.

La première écriture de l'équation de régression ne possédait pas 'd'intercept'¹⁹ et de plus le modèle était mono-factoriel.

$$\Delta Z_i = \beta_i \Delta f + \sigma_{\varepsilon i} \Delta W_1$$

2

Dans cette première équation,

- β est le coefficient directeur de la droite de régression ;
- $\sigma_{\varepsilon i} \Delta W_1$ est un aléa gaussien centré d'écart-type $\sigma_{\varepsilon i}$. Il désigne l'erreur faite sur la régression ;
- Δf décrit la variation de valeur sur l'intervalle Δt d'un indicateur économique de référence qu'on estime suivre lui-aussi un processus gaussien soit encore

$$\Delta f = \mu_f \Delta t + \sigma_f \Delta W_2$$

3

μ_f et σ_f désignent respectivement les paramètres de drifts et de volatilité propres aux variations de l'indicateur macroéconomique f .

L'absence d'ordonnée à l'origine impliquait une mauvaise qualité des régressions (exprimée à travers des coefficients R^2 faibles et des coefficients $\sigma_{\varepsilon i}$ élevés.)

Ainsi, nous avons d'abord rajouté une ordonnée à l'origine γ .

$$\Delta Z_i = \beta_i \Delta f + \gamma_i + \sigma_{\varepsilon i} \Delta W_1$$

4

Par ailleurs, en situation où la population des contreparties se répartit sur (p) pays, les variations Δf du facteur dans chaque pays suivent une distribution multivariée corrélée. Nous adopterons l'option d'une modélisation sous forme de vecteur gaussien suivant une distribution normale multivariée soit :

$$\overrightarrow{\Delta f} = N[\vec{\mu}, \Sigma]$$

5

¹⁹ Ordonnée à l'origine

Où le vecteur $\vec{\mu}$ de dimension (p) correspond aux moyennes de Δf dans chaque pays et la matrice Σ de dimension (p, p) est la matrice des covariances de Δf entre les pays. A ce sujet, on remarquera enfin qu'il est possible d'exprimer la matrice de covariance Σ en fonction du vecteur des écarts types de Δf dans chaque pays et de la matrice de corrélation R des Δf entre les pays soit :

$$\Sigma = \text{Diag}(\vec{\sigma_f}).R.\text{Diag}(\vec{\sigma_f})$$

6

Cette écriture permet d'améliorer l'explication des variations de DD par les variations de facteurs. Mais par contre, cela impose de ne devoir choisir qu'un seul et unique facteur pour toutes les contreparties. L'étude menée sur la qualification des indicateurs macroéconomiques (cf. [III.3 La sélection des indicateurs macro-économiques du modèle](#)) a montré une forte versatilité des meilleurs facteurs, en fonction du type de contrepartie (souverain, banque, ...) mais aussi des zone géographiques. Nous avons donc fait évoluer le modèle une nouvelle fois en incluant une régression linéaire multifactorielle.

Introduction de la problématique multifactorielle :

La modélisation exposée ci-dessus conduit à régresser les valeurs des ΔZ_i sur les valeurs des variations d'un facteur de référence Δf , soit une régression 'mono facteur'. Le fait que le facteur considéré connaisse une distribution multivariée selon les pays n'entre pas en considération car on suppose que les variations des ΔZ_i des contreparties sont uniquement liées au contexte de leur pays d'appartenance²⁰.

Par ailleurs, on cherche dans un modèle factoriel à maximiser sur la population des contreparties le 'poids' de la composante factorielle afin de parvenir à expliquer le plus possible des variations de notation ΔZ_i par les variations du facteur.

Aussi une première partie des travaux d'analyse factorielle va consister à qualifier parmi différents facteurs disponibles celui qui maximisera la composante factorielle des ΔZ_i . On se rend compte alors que le 'meilleur facteur' est distinct selon les contreparties. Ensuite, il s'agit de rechercher le facteur qui emporte sur la population considérée le R^2 de régression le plus important²¹. Là aussi, selon les pays et les groupes de pays, le facteur considéré n'est pas identique.

Toutefois, que l'on recherche le meilleur facteur pour une unique population répartie sur plusieurs pays ou encore que l'on adopte des facteurs différenciés par pays voire pas contrepartie, on se trouve toujours en situation de régression mono factorielle.

²⁰ Donc les régressions sont conduites au sein de chaque pays avec les valeurs de Δf propre au pays. Cette hypothèse est discutable, bien évidemment.

²¹ En pratique, il existe plusieurs manières d'opérer. On peut classer les facteurs en fonction du nombre de fois où ils sont ressortis en premier sur la population considérée, ou encore utiliser des moyennes pondérées des R^2 par les EAD, etc...

Néanmoins, une autre approche est d'expliquer les variations de distance au défaut d'une contrepartie au sein d'un pays non pas par un unique facteur, mais par plusieurs facteurs simultanément dans le cadre d'une régression multifactorielle, soit sur (m) facteurs :

$$\Delta Z_i = \sum_{k=1}^m \beta_{ik} \Delta f_k + \gamma_i + \sigma_{\varepsilon i} \Delta W_1$$

7

L'avantage d'une régression multifactorielle est que plusieurs facteurs apportent plus d'explication de la variation de distance au défaut qu'un seul. Là aussi, la question de sélectionner selon les pays et même les contreparties les facteurs les plus 'explicatifs' va se poser mais pas dans la même proportion qu'en approche mono factorielle. En effet, le choix d'un unique facteur ne se pose plus ; il faut choisir un groupe de facteurs (quand même pas trop étendu) ce qui est plus simple. En revanche, une fois ce choix opéré, la pondération pertinente entre facteurs est établie de façon optimale pour chaque contrepartie à travers les valeurs de ses coefficients de régression β_{ik} .

Dans la partie ci-après nous expliquons donc l'approche de modélisation multi factorielle.

Expression des variations de distance au défaut en 'multi factoriel'

Soit une population de n contreparties (CPs), on aura sur cette dernière des variations de distance au défaut $\Delta Z_i, i \in [1, n]$.

Ces contreparties se répartissent sur p pays. Par ailleurs, lorsque l'on s'intéressera à une contrepartie i située dans le pays j on pourra adopter la notation : $\Delta Z_{i(j)}, j \in [1, p]$.

Enfin, soit un ensemble de m facteurs qui prennent valeurs dans chacun des p pays. Dans ce cas, on pourra adopter à leur sujet la notation suivante : $\Delta f_{k(j)}$ pour parler de la variation de valeur du facteur k dans le pays j , avec l'indice k dans $[1, m]$.

En reprenant l'expression précédente mais en étant attentif au pays (j) d'appartenance de la contrepartie (i) on aura donc²² :

$$\Delta Z_{i(j)} = \sum_{k=1}^m \beta_{ik} \Delta f_{k(j)} + \gamma_i + \sigma_{\varepsilon i} \Delta W_1$$

8

Expression que l'on peut passer en notation vectorielle, soit :

$$\Delta Z_{i(j)} = \langle \vec{\beta}_i, \vec{\Delta f_{(j)}} \rangle + \gamma_i + \sigma_{\varepsilon i} \Delta W_1$$

9

²² Pour des raisons de lisibilité, nous n'avons pas généralisé l'indice (j) à tous les termes de l'expression, dès lors qu'il est bien clair qu'ils se rapportent de façon unique à la contrepartie. En fait, leur valeur dépend quand même du fait qu'on adopte en régression les seuls facteurs du pays (j) ce qui est implicite dans le modèle.

$\langle \vec{\beta}_i, \vec{\Delta f}_{(j)} \rangle$ est un scalaire qui correspond au produit scalaire du vecteur $\vec{\beta}_i$ des coefficients de régression appliqué au vecteur des variations des facteurs dans le pays j soit $\vec{\Delta f}_{(j)}$. Le paramètre γ_i est la constante de régression et $\sigma_{\epsilon i}$ est l'écart type du résidu gaussien centré de la régression multifactorielle.

Aussi, si on adopte pour décrire le vecteur $\vec{\Delta f}$ une distribution normale multivariée on aura encore l'équation (5) :

$$\vec{\Delta f} = N[\vec{\mu}, \Sigma]$$

Le vecteur $\vec{\mu} \in R^{mp}$ correspond aux m moyennes des m Δf dans chacun des p pays.

La matrice $\Sigma \in R^{mp \times mp}$ correspond à la matrice des covariances de Δf entre tous les facteurs et tous les pays.

À ce sujet, on remarquera qu'il est possible d'exprimer Σ en fonction du vecteur des écarts types $\vec{\sigma_f} \in R^{mp}$ des Δf dans chaque pays (placé dans une matrice diagonale) et de la matrice de corrélation $R \in R^{mp \times mp}$ des Δf entre les facteurs et les pays soit :

$$\Sigma = \text{Diag}(\vec{\sigma_f}) \cdot R \cdot \text{Diag}(\vec{\sigma_f})$$

10

L'équation (6) reste donc vraie.

En reprenant l'ensemble des équations on aura donc :

$$\Delta Z_i = \mu_{zi} \Delta t + \sigma_{zi} \Delta W$$

11

$$\Delta Z_{i(j)} = \langle \vec{\beta}_i, \vec{\Delta f}_{(j)} \rangle + \gamma_i + \sigma_{\epsilon i} \Delta W_1$$

12

$$\vec{\Delta f} = N[\vec{\mu}, \Sigma] \text{ avec } \Sigma = \text{Diag}(\vec{\sigma_f}) \cdot R \cdot \text{Diag}(\vec{\sigma_f})$$

13

En combinant (12) et (13) on obtient :

$$\Delta Z_{i(j)} = \langle \vec{\beta}_i, N[\vec{\mu}, \text{Diag}(\vec{\sigma_f}) \cdot R \cdot \text{Diag}(\vec{\sigma_f})] \rangle + \gamma_i + \sigma_{\epsilon i} \Delta W_1$$

14

On remarquera en combinant (11) et (12) que :

$$\begin{cases} \mu_{zi} = \gamma_i + \langle \vec{\beta}_i, \vec{\mu} \rangle \\ \sigma_{zi}^2 = [\sigma_{\epsilon i}^2] + \left[\vec{\beta}_i \cdot \text{Diag}(\vec{\sigma_f}) \cdot R \cdot \text{Diag}(\vec{\sigma_f}) \cdot \vec{\beta}_i \right] \end{cases}$$

15

Ces relations supposent que les paramètres de drift et de volatilité (μ, σ) sur les facteurs $\Delta f_{k(j)}$ et sur ΔZ_i sont calibrés sur des séries synchrones et comportant les mêmes nombres de valeurs²³.

De façon générale, on préférera s'aligner sur les paramètres μ_{zi} et σ_{zi} calibrés directement sur les contreparties, et procéder à la renormalisation des paramètres factoriels sous cette contrainte.

- **En ce qui concerne le drift μ** , on adoptera μ_{zi} au lieu et place de $\gamma_i + \langle \vec{\beta}_i, \vec{\mu} \rangle$ dans les processus de simulation.
- **En ce qui concerne la volatilité σ** , on va tout d'abord déterminer la composante idiosyncratique par recours au coefficient de détermination RSQ en utilisant la relation :

$$RSQ_{zi} = 1 - \left(\frac{\sigma_{\epsilon}}{\sigma_z} \right)^2$$

En repartant de $\sigma_{zi}^2 = [\vec{\beta}_i \cdot \text{Diag}(\vec{\sigma_f}) \cdot R \cdot \text{Diag}(\vec{\sigma_f}) \cdot {}^t \vec{\beta}_i] + [\sigma_{\epsilon}^2]$, l'expression de RSQ_{zi} conduit à $Y(0, \sigma_{\epsilon}) = \sigma_z \sqrt{1 - RSQ} \cdot Y(0, 1)$

Après élimination, on obtient la relation suivante pour la composante factorielle :

$$RSQ_{zi} \sigma_{zi}^2 = \vec{\beta}_i \cdot \text{Diag}(\vec{\sigma_f}) \cdot R \cdot \text{Diag}(\vec{\sigma_f}) \cdot \vec{\beta}_i^t$$

Nous allons rechercher à renormaliser l'expression $\vec{\beta}_i \cdot \text{Diag}(\vec{\sigma_f}) \cdot R \cdot \text{Diag}(\vec{\sigma_f}) \cdot \vec{\beta}_i^t$ sous la forme du produit d'une constante et d'une expression recourant à un vecteur lui-même normalisé. On va donc rechercher un vecteur $\vec{\alpha}$ et la constante H tels que :

$$\begin{cases} \|\vec{\alpha}\|^2 = \sum \alpha_i^2 = 1 \\ H \cdot (\vec{\alpha} \cdot R \cdot {}^t \vec{\alpha}) = {}^t (\text{diag}(\sigma_f) \beta_i) \cdot R \cdot \text{diag}(\sigma_f) \beta_i \end{cases}$$

16

Cette transformation a pour but de substituer par la suite l'expression $[RSQ_{zi} \sigma_{zi}^2] (\vec{\alpha} \cdot R \cdot \vec{\alpha}^t)$ en lieu et place de $\vec{\beta}_i \cdot \text{Diag}(\vec{\sigma_f}) \cdot R \cdot \text{Diag}(\vec{\sigma_f}) \cdot \vec{\beta}_i^t$.

Pour trouver les valeurs de $\vec{\alpha}$ et de la constante H :

$$\vec{\alpha} \cdot R \cdot \vec{\alpha}^t = \sum_{i=1}^n \alpha_i \sum_{j=1}^n R_{ij} \alpha_j \text{ soit, en remarquant que } R_{ii} = 1 \text{ ainsi que } \sum \alpha_i^2 = 1,$$

$$\vec{\alpha} \cdot R \cdot \vec{\alpha}^t = 1 + \sum_{i=1}^n \sum_{j \neq i}^n R_{ij} \alpha_i \alpha_j$$

$$H \cdot (\vec{\alpha} \cdot R \cdot \vec{\alpha}^t) = H + \sum_{i=1}^n \sum_{j \neq i}^n R_{ij} \cdot H \cdot \alpha_i \alpha_j$$

Si on adopte pour H et pour α_i les expressions suivantes :

²³ Les relations ne sont pas respectées si les paramètres factoriels $\vec{\mu_f}$ & $\vec{\sigma_f}$ sont calibrés sur une profondeur différente de celle de la notation de la contrepartie.

$$\begin{cases} H = \sum (\beta_{ik} \cdot \sigma_{fk})^2 \\ \alpha_k = \frac{\beta_{ik} \cdot \sigma_{fk}}{\sqrt{H}} = \frac{\beta_{ik} \cdot \sigma_{fk}}{\sqrt{\sum (\beta_{ik} \cdot \sigma_{fk})^2}} \end{cases}$$

17

Avec β_{ik}, σ_{fk} calibrés sur les mêmes périodes

Le vecteur α est bien normé à l'unité et est indépendant de la corrélation retenue pour R. En effet :

$$H. (\vec{\alpha} \cdot R \cdot \vec{\alpha}^t) = \sum (\sigma_f \beta_i)^2 + \sum_{i=1}^n \sum_{j \neq i}^n R_{ij} \cdot \sigma_f \beta_i \cdot \sigma_f \beta_j$$

$$H. (\vec{\alpha} \cdot R \cdot \vec{\alpha}^t) = (\text{diag}(\sigma_f) \beta_i)^t \cdot R \cdot (\text{diag}(\sigma_f) \beta_i)$$

Si on adopte la notation $R = \Lambda \cdot \Lambda^t$ fondée sur la transformation de Cholesky ^{vi} de R^{24} on a alors :

$$\Delta Z_i = \left[\gamma_i + \sum \mu_f \beta_i \right] \Delta t + Y(0, \sigma_{\epsilon i}) + \langle \vec{\beta}_i, \text{Diag}(\vec{\sigma}_f) \cdot \Lambda \cdot \vec{Y}(0,1) \rangle$$

18

où $Y(0, \sigma_{\epsilon i})$ est un aléa centré d'écart type $\sigma_{\epsilon i}$ et $\vec{Y}(0,1)$ est un vecteur gaussien (m x p) centré réduit non corrélé.

En repartant de l'expression précédente, on obtient deux expressions équivalentes²⁵ pour ΔZ_i .

La première expression s'appuie sur les seuls paramètres factoriels :

$$\Delta Z_i = \left[\gamma_i + \sum \mu_f \beta_i \right] \Delta t + Y(0, \sigma_{\epsilon i}) + \sum_{k=1}^m U_{k(j)} \beta_{ik} \sigma_{f_{k(j)}}$$

19

La seconde expression est obtenue après renormalisation sur les paramètres de la contrepartie en utilisant les relations (15), (16) et (17) :

²⁴ Le recours à la transformation de Cholesky impose que la matrice R soit semi-positive, soit encore que le nombre de variables est au moins inférieur au nombre d'observations par variables. Or, avec des séries annuelles, cette condition n'est pas respectée dès lors qu'on a affaire à un nombre de pays significatif (40 à 50) et à plusieurs facteurs.

Aussi, la matrice de corrélation obtenue avec les observations disponibles doit faire l'objet de transformations afin de rechercher la matrice définie positive la plus proche. La déformation générée par rapport à la matrice de corrélation d'origine est facile à mesurer. Ceci permet de vérifier que les écarts demeurent à un niveau acceptable, cohérent avec l'intervalle de confiance associé aux corrélations calculées sur les observés.

²⁵ S'il est nécessaire d'effectuer le calcul de \vec{U} avec la matrice de Cholesky Λ complète m X p, il est possible dans le calcul des expressions (v) et (vi) ci-dessus de se limiter aux seuls termes du pays (j).

$$\Delta Z_i = \mu_{zi} \Delta t + \sigma_{zi} \left[\sqrt{1 - RSQ_{zi}} \cdot Y(0,1) + \sqrt{RSQ_{zi}} \cdot \sum_{k=1}^m U_{k(j)} \alpha_{ik} \right]$$

20

Cette seconde expression (20) est plus adaptée aux simulations qui servent à générer les variations de distance au défaut, notamment pour passer d'aléas gaussiens à des aléas adoptant les distributions des matrices de transition.

En revanche, dans les situations où l'on souhaite appliquer des chocs factoriels, il est nécessaire de repartir de la première expression (19) pour calculer l'impact de ces derniers sur les drifts et les volatilités des contreparties : ces aspects sont développés ci-après.

Maintenant que le modèle a été entièrement décliné, une attention particulière doit être portée à deux points particuliers.

Méthode de renormalisation pour la génération des variations des Distance au défaut cohérentes avec les distributions observées sur les matrices de transition

La technique de renormalisation consiste à substituer à une variable aléatoire gaussienne corrélée X une autre variable aléatoire X' de même corrélation mais suivant une distribution F non gaussienne en utilisant la relation $X' = F^{-1}[u]$ où u est un aléa uniforme.

L'aléa uniforme est lui-même issu de l'aléa gaussien corrélé X à travers la relation $u = N[X]$. Dans ces conditions, $X' = F^{-1}[N[X]]$ suit la distribution F et reproduit la corrélation de la variable gaussienne X d'origine.

Nous générons des variations de distance au défaut $\Delta Z_i(t) = \mu_{zi} \Delta t + \sigma_{zi} \Delta w'$ calées sur les paramètres de diffusion des contreparties (μ_{zi}, σ_{zi}) mais ayant recours à $\Delta w'$ un aléa non gaussien normé centré corrélé suivant les distributions des matrices de transition.

Dans ce cas, l'aléa $\Delta w'$ est lui-même obtenu à partir de l'aléa centré corrélé Δw issu de l'expression mono factorielle : $\sqrt{1 - RSQ_{zi}} \cdot Y(0,1) + \sqrt{RSQ_{zi}} \cdot U_j(0,1)$ propre à la contrepartie concernée.

Les principes exposés pour opérer la renormalisation des aléas gaussiens issus de tirages mono-factoriels corrélés peuvent se généraliser au fonctionnement multi factoriel. A cet effet, il suffit simplement de substituer pour la valeur de Δw l'expression multi factorielle normée centrée que nous avons définie précédemment, soit :

$$\Delta w = \sqrt{1 - RSQ_i} \cdot Y(0,1) + \sqrt{RSQ_i} \cdot \sum_{k=1}^m U_{k(j)} \alpha_{ik}$$

Prise en compte des corrélations systémiques liées au risque pays

Le concept est d'introduire une dépendance entre la situation du souverain (l'Etat) et celle des contreparties au sein du pays.

La méthode la plus simple est de considérer que la notation des contreparties au sein d'un pays est plafonnée à celle du souverain. En fait, cette option est fortement discutable, car une situation de défaut du souverain n'entraîne pas forcément celle des acteurs économiques du pays concerné, bien qu'il soit évident que ces derniers sont néanmoins fortement impactés.

Pour cette raison, nous avons adopté une autre approche qui consiste à introduire une corrélation plus ou moins forte entre la notation du souverain et celle des contreparties du pays. L'idée est de considérer que plus la situation du pays se dégrade, plus cette corrélation trouve à s'exprimer.

L'aléa normé centré pour l'Etat du pays **ASOV** réutilisé pour tous les tiers du pays répondra à l'expression :

$$ASOV = \sqrt{1 - RSQ_{Sov}} \cdot Y(0,1) + \sqrt{RSQ_{Sov}} \cdot \sum_{k=1}^m U_{k(j)} \alpha_{Sov k}$$

En fonction de cet aléa, on détermine le niveau de notation de l'Etat du pays et, en conséquence, la valeur du coefficient de corrélation systémique **ρ_s des contreparties à l'Etat du pays (décliné le cas échéant par secteurs)**.

Ensuite pour les tiers (i) du même pays, on va calculer leurs variations de DD :

$$\Delta Z_i(t) = \mu_{zi} + \sigma_{zi} \cdot [\rho_s ASOV + \sqrt{1 - \rho_s^2} \cdot [\sqrt{1 - RSQ_i} \cdot Y(0,1) + \sqrt{RSQ_i} \cdot \sum_{k=1}^m U_{k(j)} \alpha_{i k}]]$$

Il s'agit là des notations pour des aléas gaussiens. En tout état de cause, les procédures de renormalisation décrites précédemment pour se caler sur les distributions des matrices de transitions sont à appliquer à cet aléa gaussien.

3. La sélection des indicateurs macro-économiques du modèle

a. Choix du panel d'indicateurs macroéconomiques

Initialement, une cinquantaine d'indicateurs a été recueillie. Ils sont majoritairement issus de la Banque Mondiale mais aussi de MSCI (Morgan Stanley Capital International) et de Eurostat. Ils ont été choisis car ils semblaient tous pertinents^{vii} - d'un point de vue économique - pour expliquer des variations de distance au défaut.

Le nombre d'indicateurs a finalement été réduit à 17 après l'étude de la profondeur des historiques disponibles. Les proportions de données disponibles sur 50 ans (1969-2018), 40 ans (1979-2018), 30 ans (1989-2018) et 20 ans (1999-2018) ans ont été calculées. Les indicateurs pour lesquels trop peu de données étaient disponibles, particulièrement pour les tranches 20 ans et 30 ans ont été éliminés. Parmi les indicateurs finalement écartés pour cette raison il y a par exemple le Classement de la qualité de la gestion budgétaire et financière ou les stocks de la dette extérieure en dollars américains courants.

Après la recherche documentaire, sur les techniques de ratings ([cf. infra](#)), 3 indicateurs ont été rajoutés. Le PIB/ Habitant en \$ US Courants issus des de la Banque Mondiale ainsi que les taux de changes et les taux 3 mois (ces deux derniers étant imposés par la réglementation), tous deux issus de Bloomberg. Portant donc le nombre d'indicateurs à 20.

Remarquons également qu'il existe une pléthore de sources de données encore exploitable, notamment la série de 200 indicateurs 'Doing Business' de la Banque Mondiale.

Voici les indicateurs utilisés :

- 1) **Accumulation nette de passifs (% du PIB):** Part nette des 'engagements' et dettes du gouvernement.
- 2) **Balance des paiements courants (% du PIB):** Somme des exportations nettes de biens, de services, de revenus nets et de transferts nets courants.
- 3) **Capacité (+)/besoin (-) de financement (% du PIB) du secteur public :** Mesure des ressources financières mises à la disposition ou au contraire prises par le secteur public sur les autres secteurs de l'économie ou l'étranger.
- 4) **Capitalisation boursière des entreprises intérieures cotées (% du PIB):** Valeur de marché pour les sociétés nationales cotées en bourse.
- 5) **Chômage:** Part de la population active qui est sans travail mais disponible et à la recherche d'un emploi.
- 6) **Créances sur le gouvernement central (% du PIB):** Créances sur le gouvernement central (SFI ligne 52AN ou 32AN).
- 7) **Croissance du PIB (Δ PIB % annuel):** Pourcentage annuel de croissance du PIB aux prix du marché basé sur les devises locales constantes.

- 8) **Croissance du PIB par habitant:** Taux annuel de croissance du PIB par habitant basé sur les devises locales constantes.
- 9) **Dépenses nationales brutes (% du PIB)** Somme des dépenses en consommation finale des ménages, des dépenses en consommation finale du gouvernement et de la formation brute de capital. Les données sont en dollars américains courants.
- 10) **Épargne intérieure brute (% du PIB):** PIB moins les dépenses en consommation finale.
- 11) **Exportations de biens et de services (% du PIB):** Valeur de tous les biens et autres services offerts au reste du monde.
- 12) **Formation brute de capital (% du PIB):** Montants de variation du poste des immobilisations corporelles de l'économie plus les variations nettes des inventaires.
- 13) **Importations de biens et de services (% du PIB):** Valeur de tous les biens et autres services reçus du reste du monde.
- 14) **Inflation, déflateur du PIB (% annuel):** L'inflation mesurée par le taux annuel de croissance du déflateur implicite du PIB.
- 15) **PIB (\$ US constants de 2010):** Somme de la valeur ajoutée brute de tous les producteurs résidents, plus toutes les taxes sur les produits et moins les subventions non incluses dans la valeur des produits, sans déductions. Les montants en dollars pour le PIB sont convertis à partir des devises locales en utilisant les taux de change officiels de 2010.
- 16) **PIB (\$ US courants):** Les montants en dollars pour le PIB sont convertis à partir des devises locales en utilisant les taux de change officiels de l'année considérée.
- 17) **PIB par habitant (\$ US courants):** C'est le PIB en US Courants divisé par la population en milieu d'année.
- 18) **Taux de changes:** Pour les pays en Euro, il s'agit du rapport du prix de l'Euro sur le prix du Dollar US, tandis que pour tous les autres pays, il s'agit du rapport de la devise locale avec l'Euro.
- 19) **Taux 3 mois :** Ce sont pour la zone Euro, les Euribor 3 mois et pour les autres zones géographiques, les taux 3 mois, de la Banque Centrale Compétente
- 20) **Total des réserves (comprend l'or, \$ US courants):** Positions d'or monétaire, de droits de tirage spéciaux, de réserves des membres du FMI détenues par le FMI et de positions de devises étrangères placées sous le contrôle des autorités monétaires.

Homogénéisation des indicateurs:

Afin de pouvoir effectuer des comparaisons cohérentes, il a fallu harmoniser les données. Nous avons choisis que les données soient homogènes à des log-rendements $\left(\frac{f_{ti}-f_{ti-1}}{f_{ti-1}}\right)$ ou à des rapports de premier ordre.

En conséquence, selon la nature de l'indicateur envisagé, nous avons regardé s'il pouvait être utilisé directement ou si des transformations étaient à opérer. Les règles retenues sont les suivantes :

Type d'indicateur	Transformation opérée
Variation de type log rendement en % (exemple : Δ PIB %)	Aucune
Ratio déjà exprimés en % (exemple : taux de chômage, ratios exprimés en % du PIB du genre dette/ PIB ...)	Variation entre les deux valeurs annuelles soit: $\left(\frac{e_{ti}}{f_{ti}}\right) - \left(\frac{e_{ti-1}}{f_{ti-1}}\right)$
Montant ou quantité	Calcul d'un log-rendement $\left(\frac{f_{ti}-f_{ti-1}}{f_{ti-1}}\right)$

Tableau 5: Homogénéisation des indicateurs Tableau 1

	Expression Reçue	Transformation effectuée
Accumulation nette de passifs	Exprimé en % du PIB par année	variations
Balance des paiements courants	Exprimé en % du PIB par année	variations
Capacité (+)/besoin (-) de financement	Exprimé en % du PIB par année	variations
Capitalisation boursière des entreprises intérieures cotées	Exprimé en % du PIB par année	variations
Chômage	Exprimé en % Population total	variations
Créances sur le gouvernement central	Exprimé en % du PIB par année	variations
Croissance du PIB (Δ PIB%)	Log rendement	Aucune

Croissance du PIB par habitant	Log rendement	Aucune
Dépenses nationales brutes	Exprimé en % du PIB par année	variations
Épargne intérieure brute	Exprimé en % du PIB par année	variations
Exportations de biens et de services	Exprimé en % du PIB par année	variations
Formation brute de capital	Exprimé en % du PIB par année	variations
Importations de biens et de services	Exprimé en % du PIB par année	variations
Inflation, déflateur du PIB (% annuel)	% Pourcentage PIB de référence	variations
PIB (\$ US constants de 2010)	Montant	Calcul log-rendements
PIB (\$ US courants)	Montant	Calcul log-rendements
PIB par habitant (\$ US courants)	Montant	Calcul log-rendements
Taux de Changes	Rapport	Calcul log-rendements
Taux 3 Mous	Rapport	Calcul log-rendements
Total des réserves (comprend l'or, \$ US courants)	Montant	Calcul log-rendements

Tableau 6: Homogénéisation des indicateurs Tab 2

b. Détermination du meilleur indicateur parmi le panel en mono-factoriel

Ainsi qu'il a été dit, le modèle était initialement mono-factoriel. Alors que le panel d'indicateurs est choisi, il convient à présent de choisir le meilleur indicateur parmi ce panel.

Dans cette partie, lorsque le modèle était toujours mono factoriel, plusieurs définitions ont été attribuées au terme « meilleur indicateur ». Plusieurs champs de classification ont en conséquence été étudiés

Meilleur indicateur par maximisation du coefficient R^2 de régression

Ainsi qu'il sera développé en [annexe](#), une régression sera d'autant plus correcte que le coefficient R^2 sera proche de 1. Ainsi, afin de choisir le meilleur indicateur, nous avons voulu classer les indicateurs en fonction de cet aspect.

Nous avons défini, 2 critères.

Le premier est le plus naturel, il compte tout simplement, après avoir testé toutes les indicateurs macroéconomiques sur chacune des contreparties, combien de fois est-ce que chacun des indicateurs à le plus grand R^2 .

Le second consiste à considérer le vecteur de R^2 (de taille le nombre de contreparties de chacun des facteurs) puis à calculer la proportion de R^2 supérieure à une certaine valeur (les résultats à 0.2 et 0.5 sont présentés dans le, tableau suivant.

	Nb X Meilleur Indic	% $R^2 >$ 0.2	% $R^2 >$ 0.5
Croissance du PIB	3	22%	2%
Accumulation nette de passifs	8	18.75%	6.25%
Balance des paiements courants	0	10%	2%
Capacité /besoin de financement	4	8.57%	2.86%
Capitalisation boursière	19	24%	8%
Chômage	11	5.26%	0%
Créances sur le gouvernement central	3	16.67%	0%
Croissance du PIB / HBT	6	24%	2%
Dépenses nationales brutes	9	14%	4%
Épargne intérieure brute	17	24%	6%
Exportations de biens et de services	18	29.09%	7.27%
Formation brute de capital	4	12%	4%
Importations de biens et de services	7	14.55%	5.45%
Inflation, déflateur du PIB	19	18%	4%
PIB \$ US constants	0	22%	2%
PIB \$ US courants	8	6%	2%
Total des réserves	9	16%	6%

Tableau 7: Performance des indicateurs en mono factoriel Tab 1

En plus du tableau nous pouvons également visualiser les histogrammes des coefficients de corrélations.

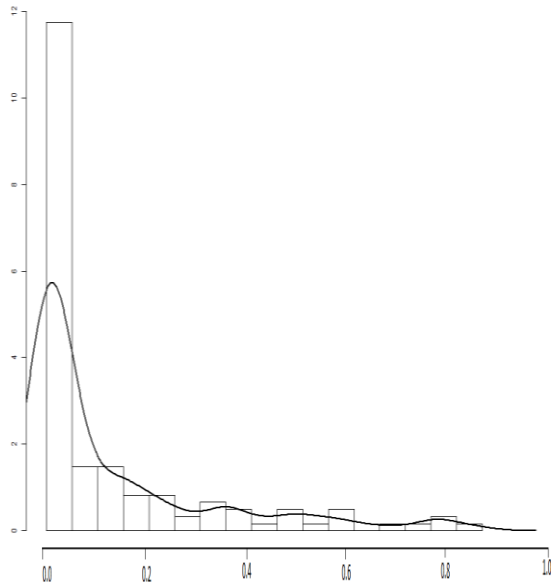


Figure 5: Histogramme des R^2 : Accumulation nettes de passifs

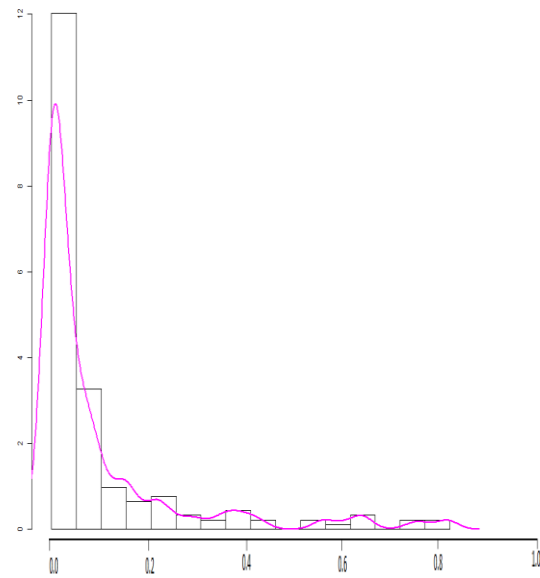


Figure 4: Histogramme des R^2 : Inflation / Déflateur du PIB

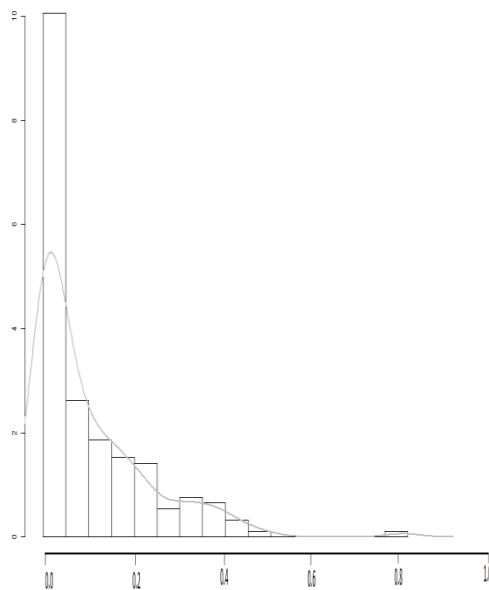


Figure 7: Histogramme des R^2 : Croissance du PIB / Hbt

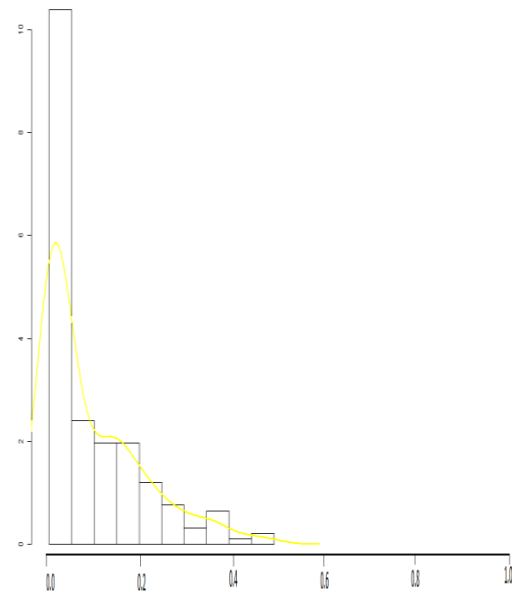


Figure 6: Histogramme des R^2 : Epargne Intérieure brute

Des résultats précédents on observe que les Exportations de Biens et Services ainsi que les Exportations se distinguent pour ce type de qualification.

Analyse des corrélations inter-pays à indicateurs constants

Le modèle est basé sur la corrélation qu'il existe entre les tiers d'un pays et leur souverain et entre les corrélations qu'il existe en les pays, notamment pour les [chocs factoriels](#).

Un indicateur qui permet de maximiser les corrélations inter-pays porterait donc un grand intérêt. Nous décidons donc de les classer selon ce critère également.

Nous calculons donc des matrices de corrélation linéaires des séries temporelles de variations de facteurs entre pays. Sur les matrices obtenues, nous retenons la matrice triangulaire inférieure diminuée de la diagonale. Nous en déduisons les indicateurs de mesure normatifs suivants calculés sur les valeurs absolues des corrélations $|\rho|$ inter-pays, à facteurs constants :

- Nombre d'occurrences de coefficients 'forts' ($|\rho| > 0.7$) ;
- Moyenne des coefficients ;
- Ecart type des coefficients ;
- Distributions des coefficients de corrélations pour chaque indicateur.

	<u>Moyennes</u>	<u>Ecart-Type</u>	<u>%coefs > 0.7</u>
Croissance du PIB	0.303	0.214	6.58%
Accumulation nette de passifs	0.226	0.180	5.37%
Balance des paiements courants	0.164	0.127	2.56%
Capacité (+)/besoin (-) de financement	0.269	0.178	4.50%
Capitalisation boursière	0.526	0.208	24.28%
Chômage	0.219	0.389	9.88%
Créances sur le gouv. Central	0.174	0.146	3.52%
Croissance du PIB par habitant	0.156	0.126	2.17%
Dépenses nationales brutes	0.203	0.177	3.92%
Épargne intérieure brute	0.284	0.207	5.95%
Exportations	0.224	0.168	3.85%
Formation brute de capital	0.325	0.223	8.94%
Importations	0.164	0.166	3.10%
Inflation, déflateur du PIB	0.298	0.216	7.13%
PIB (\$ US constants de 2010)	0.359	0.216	9.65%
PIB (\$ US courants)	0.358	0.248	14.06%
Total des réserves (\$ US courants)	0.209	0.185	4.81%

De même que pour l'analyse des coefficients R^2 , les histogrammes permettent de bien visualiser les résultats.

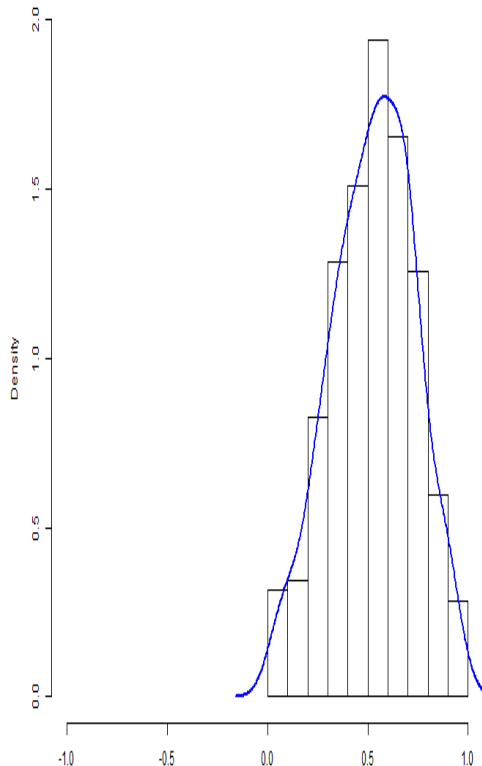


Figure 9: Histogramme coefficients de corrélation : Capitalisation boursière

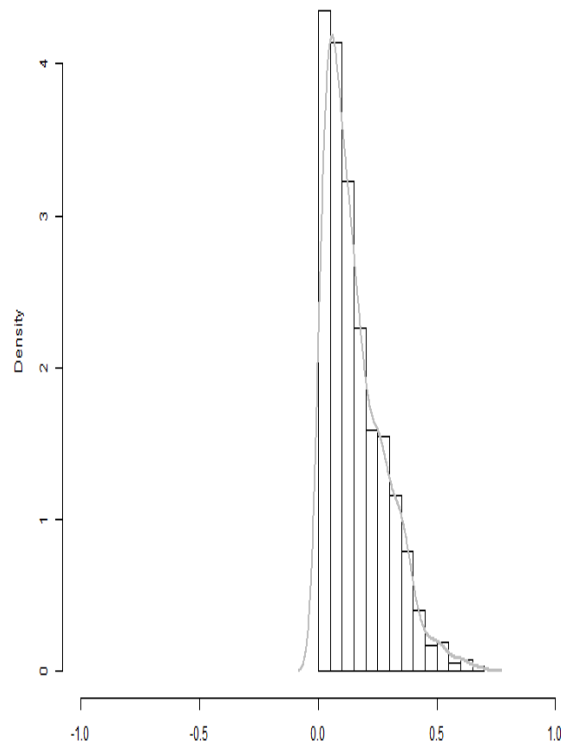


Figure 8: Histogramme coefficients de corrélation: Dépenses nationales brutes

Les Capitalisation boursières et le PIB en US Courants semblent se distinguer d'après ce critère.

Distribution des indicateurs : qualification des lois marginales

Le dernier axe de classification est la reproductibilité des comportements des distributions afin d'avoir des simulations suffisamment proches de la réalité.

En effet, les modèles factoriels reposent sur l'hypothèse que les variations temporelles des facteurs suivent des comportements de lois normales (ou log-normales). Nous constituons les échantillonnages des variations des facteurs à partir des séries temporelles pour analyser la forme des distributions. Nous partons d'une hypothèse de normalité (paramètres de drift et de volatilité μ , σ) et nous testons la validité de cette dernière à travers des tests classiques (Shapiro-Wilk). Lorsque ces tests ne sont pas probants, nous utilisons la loi [NIG](#) qui outre un drift et une volatilité embarque une skewness et un kurtosis (respectivement moments d'ordre 3 et 4) et nous estimons

les paramètres par la méthode de maximisation de vraisemblance. Nous bouclons l'analyse pour valider le taux de 'fitting' obtenu par des tests de type Q-Q plot contre la NIG estimée et un test de Kolmogorov-Smirnov.

Rappelons l'interprétation des p-valeurs :

- p-valeur < 0.01 très forte présomption contre H_0 ;
- $0.01 < \text{p-valeur} < 0.05$ présomption contre H_0 ;
- $0.05 < \text{p-valeur} < 0.1$ faible présomption contre H_0 ;
- p-valeur > 0.1 aucune présomption contre H_0 .

Dans le cas d'un test de normalité (Shapiro-Wilk) H_0 est que la distribution est normale. Et pour le test de Kolmogorov-Smirnov pour la NIG, H_0 est que l'échantillon suit la NIG.

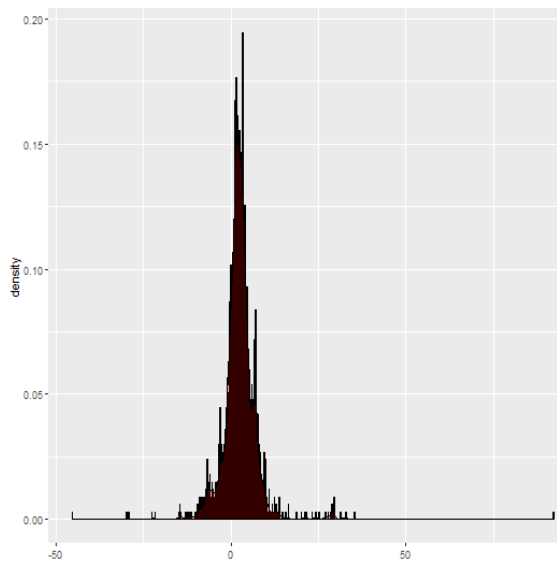


Figure 11: Histogramme de la croissance du PIB/Hbt avec la NIG estimée

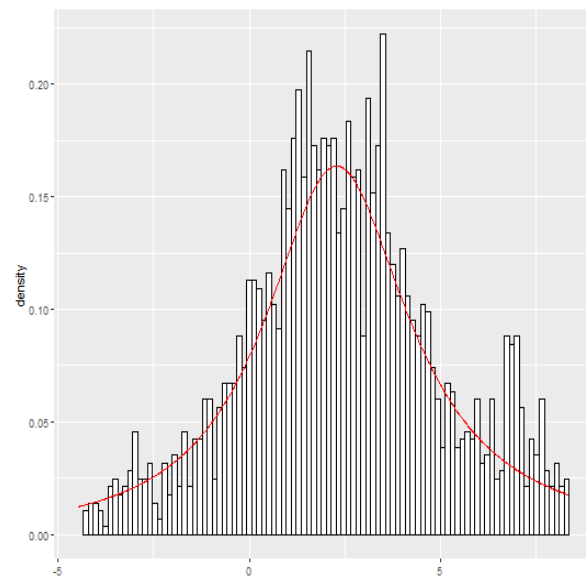


Figure 10: Histogramme de la croissance du PIB/Hbt

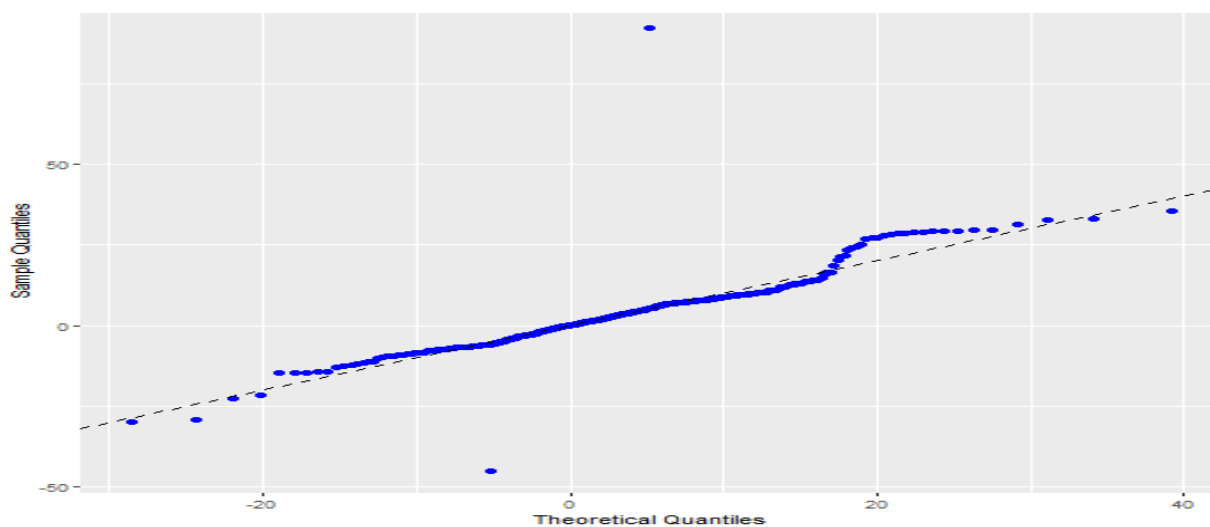


Figure 12: QQ-Plot contre la NIG estimée

	p-value Test de Normalité	p-value Test de la NIG
Croissance du PIB	0	0.658
Accumulation nette de passifs	0	0
Balance des paiements courants	0	0
Capacité/besoin fincmnt.	0	0.022
Capitalisation boursière	0	0,000739
Chômage	0	0.006
Créances sur le gouv. Central	0	0.0149
Croissance PIB/HBT	0	0.369
Dépenses nationales brutes	0	0,00034
Épargne intérieure brute	0	0.292
Exportations	0	0,000537
Formation brute de capital	0	0.139
Importations	0	0,00147
Inflation, déflateur du PIB	0	0,00351
PIB US constants	0	0.067
PIB US courants	0	0.083
Total des réserves	0	0.034

Tableau 9: Analyse de la qualité de fitting des distributions

On constate tout d'abord qu'il est clair que les distributions des indicateurs ne sont pas normales du tout. Ensuite, certains indicateurs sont particulièrement bien fittés par la NIG et leur comportement devrait en conséquence pouvoir être bien simulé. Il s'agit notamment de la croissance du PIB et de l'Épargne Intérieure Brute.

Nécessité d'inclure plusieurs facteurs :

Les 3 axes de qualifications envisagés ont résulté à des indicateurs macro-économiques différents. Dès cette conclusion, nous sentons la nécessité d'inclure plusieurs indicateurs.

Plutôt que de passer directement à un modèle multifactoriel, ce qui implique de grands changements dans le modèle mathématique, nous essayons à partir des indicateurs présents de construire un indicateur composite, construit comme une combinaison linéaire d'un certain nombre d'indicateurs. Remarquons que si cette solution avait été adoptée, il aurait été nécessaire de donner un sens économique à cet indicateur macroéconomique.

Deux techniques ont été testées.

Première technique :

Cette technique cherche à déterminer un nombre d'indicateurs qui seraient le plus important et qui n'entraîneraient pas de redondance et de trouver une combinaison linéaire de ces indicateurs qui pourrait améliorer la qualité des régressions au sens des R^2 .

Nous procédons à une [Analyse en Composante Principale](#) (ACP). Nous en déduisons donc les dimensions principales ainsi que les contributions de chacun des indicateurs macroéconomiques à ces composantes principales.

Nous procédons ensuite à un [Clustering](#) sur les indicateurs afin de déterminer les indicateurs ayant les mêmes caractéristiques et qui donc pourraient être redondants.

Enfin, dans chacun des clusters, nous sélectionnant l'élément dont la norme euclidienne des contributions aux dimensions principales de la PCA est la plus grande. Nous avons ensuite pris comme poids, la norme des contributions aux dimensions principales normalisées

$$\gamma_i = \frac{\| \text{Contribution} \|_{i_2}}{\sum_k \| \text{Contribution} \|_{k_2}}$$

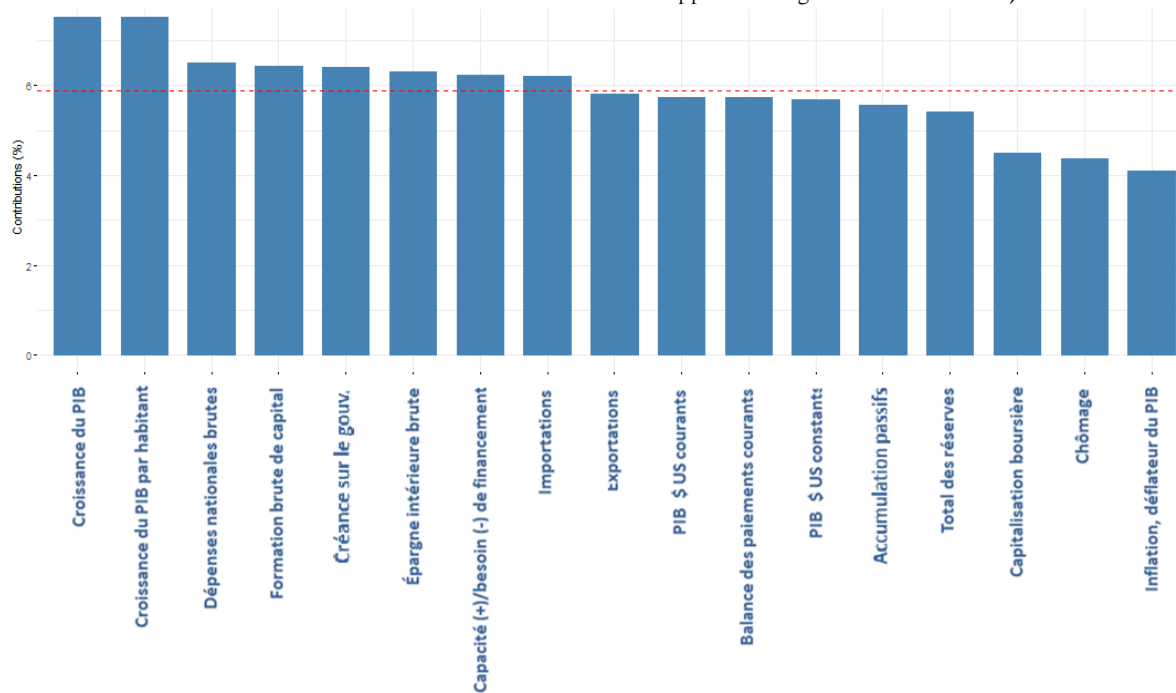


Figure 13: Contribution des Indicateurs aux 4 dimensions principales:

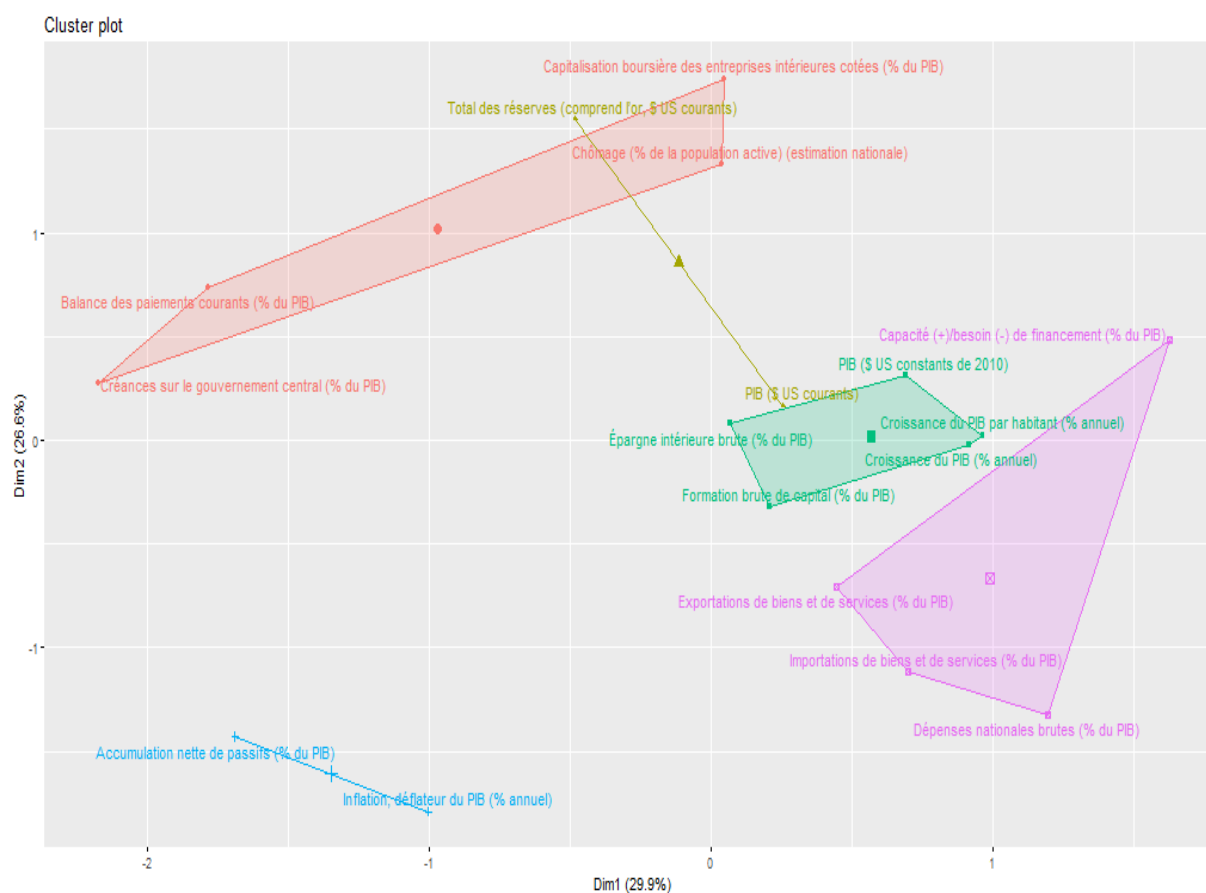


Figure 14: Clustering sur les indicateurs macroéconomiques

Deuxième technique :

Celle-ci est beaucoup plus intuitive. Nous avons corrélé chaque indicateur aux variations de distance au défaut. Nous en avons déduits les 4 indicateurs maximisant cette corrélation ρ . Les corrélations normalisées ont été utilisées comme poids. $\gamma = \frac{\rho}{\sum \rho}$.

Résultats :

Malheureusement, aucune de ces deux techniques ne s'est avérée être pertinente. Il n'y a pas une amélioration - au sens du R^2 - significative.

	1 ^{ère} technique	2 ^{ème} technique
Proportion améliorée	14.89%	23.40%
Moyenne des différences (Modifié – Classique)	-25.72%	-13.15%

Tableau 10: Approche particulières Clustering et PCA

D'abord, la versatilité des meilleurs indicateurs en fonction du critère appliqué, ensuite, la difficulté de construire des indicateurs composites efficaces et enfin les résultats présentés dans le tableau 11 (qui compte par pays le nombre de fois où un indicateur a été classé comme meilleur) et la figure 17 (qui mesure l'évolution de l'effet factoriel en fonction du nombre d'indicateurs, mettant en exergue les grandes variations qu'il y'a d'une zone géographique à une autre pour les

meilleurs indicateurs au sens du coefficient R^2), ont définitivement annihilé la possibilité de rester en monfactoriel

	Total des réserves	PIB (\$ US courants)	Balance des paiements courants	Inflation, déflateur du PIB (% annuel)	Dépenses nationales brutes (% du PIB)	Accumulation nette de passifs (% du PIB)	Capacité (+)/besoin (-) de financement	Croissance du PIB (% annuel)	Épargne intérieure brute (% du PIB)	Croissance du PIB par habitant (% annuel)	Importations	Capitalisation boursière	Créances sur le gouvernement central	Formation brute de capital (% du PIB)	PIB (\$ US constants de 2010)	Exportations	Chômage
Europe du Nord	4	4	3	0	1	0	0	2	0	1	0	2	1	1	1	1	0
Pays Germaniques	11	9	9	7	4	0	2	3	1	3	2	0	0	4	2	0	0
Benelux	2	6	0	2	2	1	1	1	2	1	3	0	0	0	0	0	1
France	6	12	8	0	3	0	2	1	1	0	1	3	0	1	3	1	3
Espagne	23	3	7	3	0	9	4	0	0	0	1	0	0	2	0	0	0
Italie	8	3	1	17	1	0	0	0	1	0	0	0	0	0	0	0	0
Europe du Sud	0	4	4	3	1	0	0	1	0	1	1	0	0	0	1	0	0
Europe du Centre	4	5	5	4	4	3	7	2	3	3	3	2	3	2	0	0	0
Europe de l'Est	0	3	0	1	0	0	0	1	0	1	0	1	0	0	0	0	0
Turquie	4	1	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0
Zone Royaume Unis	1	2	7	4	24	6	5	2	2	1	0	5	3	0	1	1	1
Amerique du Nord	5	3	3	2	1	13	0	4	9	2	1	1	6	0	2	3	0
Asie Pacifique	5	4	4	3	1	2	1	4	0	5	3	0	0	2	0	0	1
Nb. Occurrences	73	59	52	46	42	34	22	21	19	18	15	14	13	12	10	7	6

Tableau 11: Meilleurs Indicateurs par Zone Géographique

Le tableau précédent a permis de hiérarchiser les facteurs en fonction de leur importance pour les différentes zones géographiques.

Afin de mieux visualiser l'effet de conserver ou non plusieurs indicateurs, nous avons tracé pour chaque groupe pays la courbe de la proportion d'effet factoriel cumulée expliqué.

Autrement dit, pour chaque groupe de pays nous avons calculé la proportion de R^2 de qualité optimale obtenue en injectant au fur et à mesure les indicateurs dans l'ordre du tableau précédent.

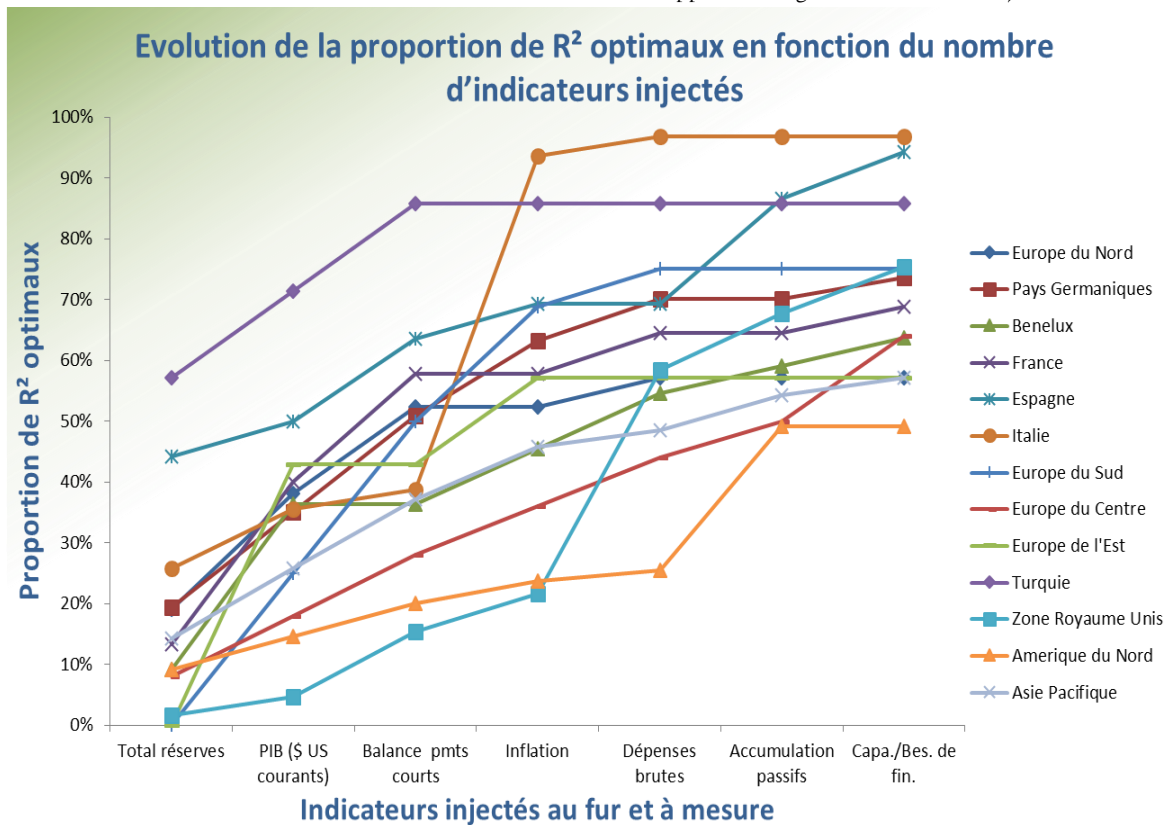


Figure 15 : Evolution proportion R^2 optimaux fonction nombre d'indicateurs injectés:

Un autre résultat fondamental qui se voit grâce aux courbes est que la hiérarchie parmi les facteurs est très variable d'un groupe à un autre. On constate en effet par exemple, que pour la Zone Royaume-Unis, lorsque l'on a injecté (le Total des réserves, le PIB \$Courants et la balance des paiements courants, nous sommes encore à 20% de R^2 optimaux alors que pour les Pays Germaniques, nous sommes déjà à 50%. Pourtant lorsque les 7 indicateurs sont injectés, tous les deux groupes sont à 74%.

Il est donc en effet, nécessaire de passer à un modèle multifactoriel

c. Détermination de la meilleure famille d'indicateurs en multifactoriel

La sélection des indicateurs pour le modèle multifactoriel s'est basée sur un seul et unique axe qui est celui de la maximisation des coefficients ajustés $R^2_{ajusté}$. Ainsi, dans tout ce qui suit, l'expression "meilleur indicateur" désignera "qui maximise le coefficient de détermination ajusté $R^2_{ajusté}$ ".

La première approche que nous avons était très naïve. Nous avons mis en place un "greedy algorithm"²⁶ afin de déterminer pour chacune de nos contreparties, la meilleure combinaison possible parmi les indicateurs ayant au plus m indicateurs

- On décide de garder pour chaque contrepartie seulement **m facteurs au plus** (dans nos travaux nous prenons $m = 5$ facteurs).
- Pour chacune des contreparties **on effectue toute les régressions multifactorielles possibles utilisant au moins 1 facteur et au plus 5 facteurs** parmi les 19²⁷. Cela implique donc pour chaque contrepartie de faire 16 663 régressions²⁸
- Enfin, on **retient** pour cette contrepartie **le groupe de facteurs (il contient donc 1 à 5 éléments) qui maximise son R^2 ajusté**.

Ci-dessus un exemple simple sur 3 facteurs a, b et c et sur une contrepartie.

Facteurs utilisés pour la régression	R^2_{adj}	Facteurs Retenu
a, b et c	0%	
a et b	43%	a et b
a et c	37%	
a	17%	
b	1%	
c	17%	

Tableau 12: Exemple de trace de du greedy algorithm

²⁶ Algorithme glouton

²⁷ Les taux 3 mois n'avaient pas encore été inclus lors de ces travaux

²⁸ Nb $k - \text{uplets}, (k \leq 5) = \sum_{k=0}^5 C_{19}^k = 16\,664$ on retire le k -uplet "vide" d'où les 16 663 régressions

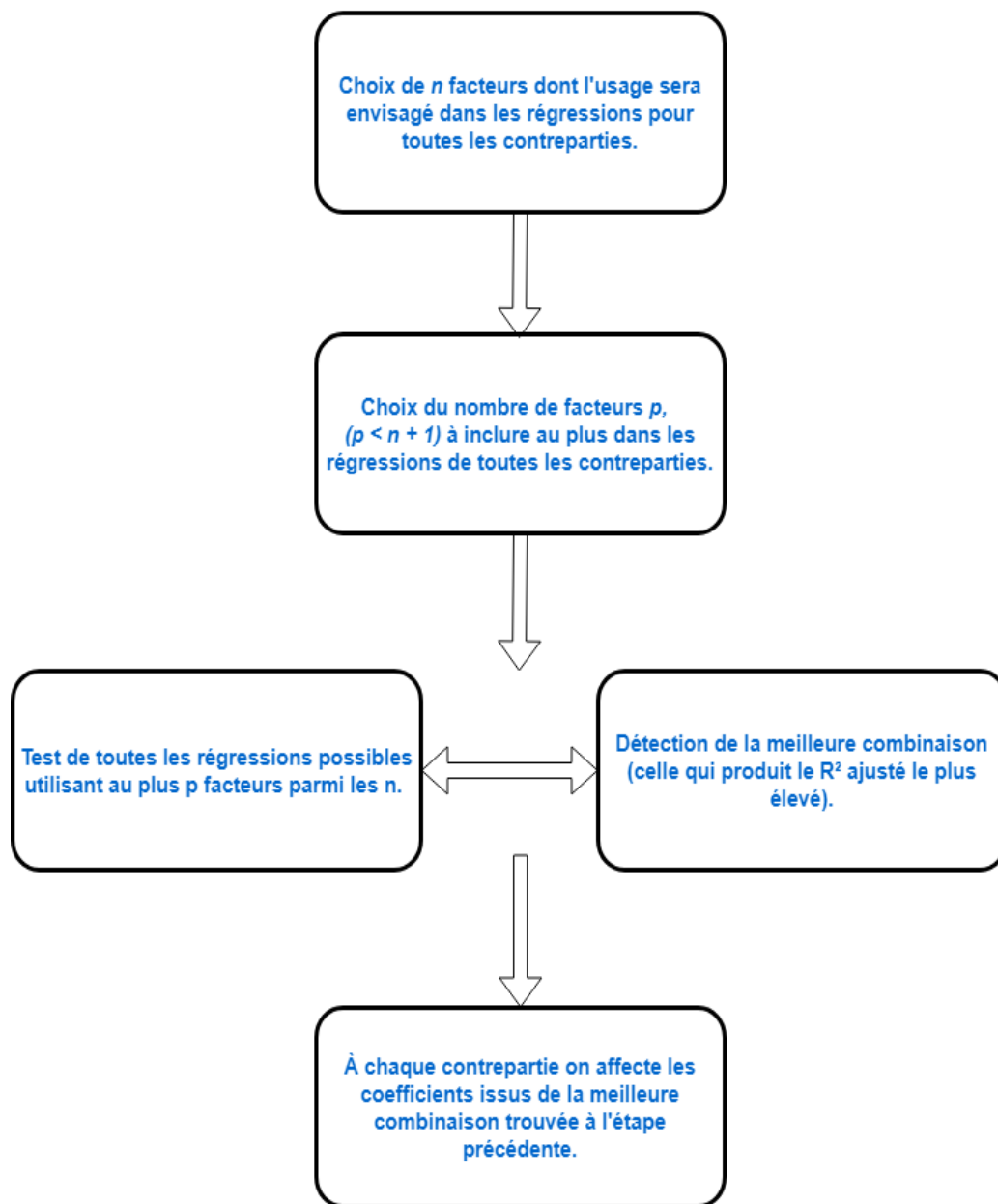
Greedy Alorithm

Figure 16: Greedy Algorithm:

L'idée était qu'éventuellement **un groupe de 5 facteurs se distinguerait** et qu'il serait donc pertinent de garder ce groupe de 5 indicateurs macroéconomiques. Malheureusement, ainsi que le montre le tableau ci-dessus, cela n'a pas été le cas.

Au moins 16 des 19 facteurs sont présents de manières quasi équivalentes particulièrement dans les quintuplets (avant dernière colonne du tableau).

En effet, en simulation on ne peut pas retenir 16 facteurs. Il faudrait manipuler des matrices de corrélations de tailles $(\text{Nb pays} \times 16) \times (\text{Nb pays} \times 16)$ sachant que le nombre de pays est autour de 45. De plus, économiquement ce ne serait pas très sensé.

	Number of Indicators					Total
	1	2	3	4	5	
Balance paiements courants	1	0	1	4	151	157
Croissance PIB/Habitant	1	1	2	7	146	157
Capacite / Besoin de Financement	0	2	2	5	141	150
Inflation	0	0	2	9	137	148
Croissance PIB	0	0	2	10	132	144
Total des réserves	0	2	0	5	137	144
Taux de Changes	1	1	2	7	132	143
Depenses Nationales Brutes	1	1	1	9	124	136
PIB/habitant US courants	0	2	0	9	113	124
Formation Brute de Capital	0	0	2	4	115	121
Capitalisation Boursière	1	0	2	7	106	116
PIB US Courants	0	1	0	4	111	116
Importations	0	1	0	11	100	112
Exportations	0	2	1	8	100	111
PIB US Constants	0	1	1	7	102	111
Epargne Intérieure Brute	0	0	0	2	103	105
Accumulation Nette de Passifs	0	1	3	2	92	98
Creances sur le Gouvernement	0	0	3	1	74	78
Chômage	0	1	0	1	29	31

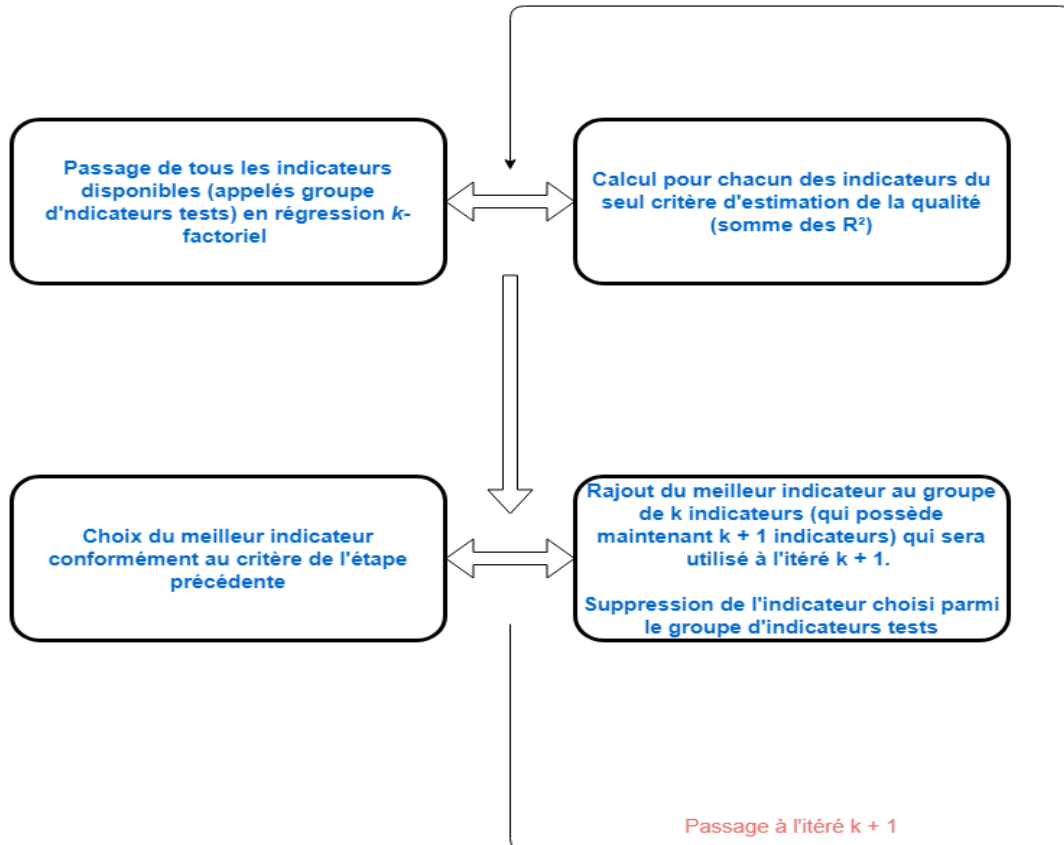
Tableau 13: Occurrences des indicateurs dans les k-uplets résultants du greedy algorithm

Remarquons néanmoins que même si les résultats obtenus ne sont pas exploitables pour le choix d'indicateurs, ils sont très utiles pour l'analyse globale de la population des contreparties. En effet, ils servent de référence. Nous obtenons une mesure de la qualité maximale de régression que l'on peut obtenir en considérant des groupes d'au plus 5 indicateurs parmi notre panel d'indicateurs (sans transformation additionnelle, incluant notamment des lags). La moyenne des R^2 ajustés est de 81.91% et celle des R^2 classiques est de 88.12%.

Le greedy algorithm permet au mieux de réduire à 16 indicateurs les 19 : ceci ne permet pas de déboucher sur des fonctionnements acceptables en simulation. Nous avons mis en place un algorithme menant nécessairement à 5 indicateurs identiques sur toute la population.

Cet algorithme utilise un autre algorithme de sélection des indicateurs appelé "ascending algorithm"²⁹. En fonction d'un critère, il choisit un indicateur à rajouter à la liste des indicateurs utilisables pour régression. Dans nos travaux le critère était la maximisation de la moyenne des R^2 .

Itération k de l'ascending algorithm



Premier itéré de l'algorithme



Dernier itéré n de l'algorithme

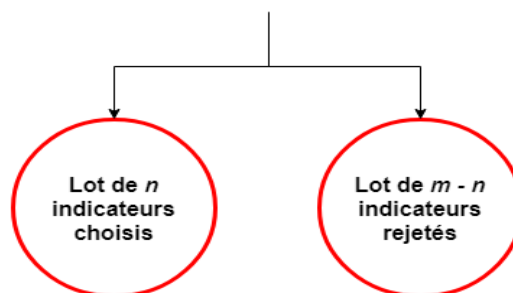


Figure 17: Ascending algorithm

²⁹ Algorithme ascendant

L'algorithme finalement utilisé se décrit comme suit :

À chaque itération, nous utilisons, un “ascending algorithm”³⁰ afin de sélectionner un indicateur à ajouter au lot d'indicateurs disponibles pour la régression. À l'itéré 1 nous aurons donc choisi un indicateur parmi le panel de 20 indicateurs, puis à l'itéré suivant nous aurons rajouté un nouvel indicateur au lot (il contient donc 2 indicateurs) et ainsi de suite.

À chaque fois qu'un indicateur est rajouté au lot, un “greedy algorithm” est appliqué sur l'ensemble des contreparties afin d'évaluer l'ensemble des combinaisons de facteurs possibles afin de choisir pour chacune de ces contreparties la combinaison optimale possible en ne se limitant qu'aux facteurs présents dans le lot.

Ainsi, à chaque itéré nous mesurons par plusieurs paramètres l'effet factoriel optimal obtenu.

Une Itération de l'algorithme

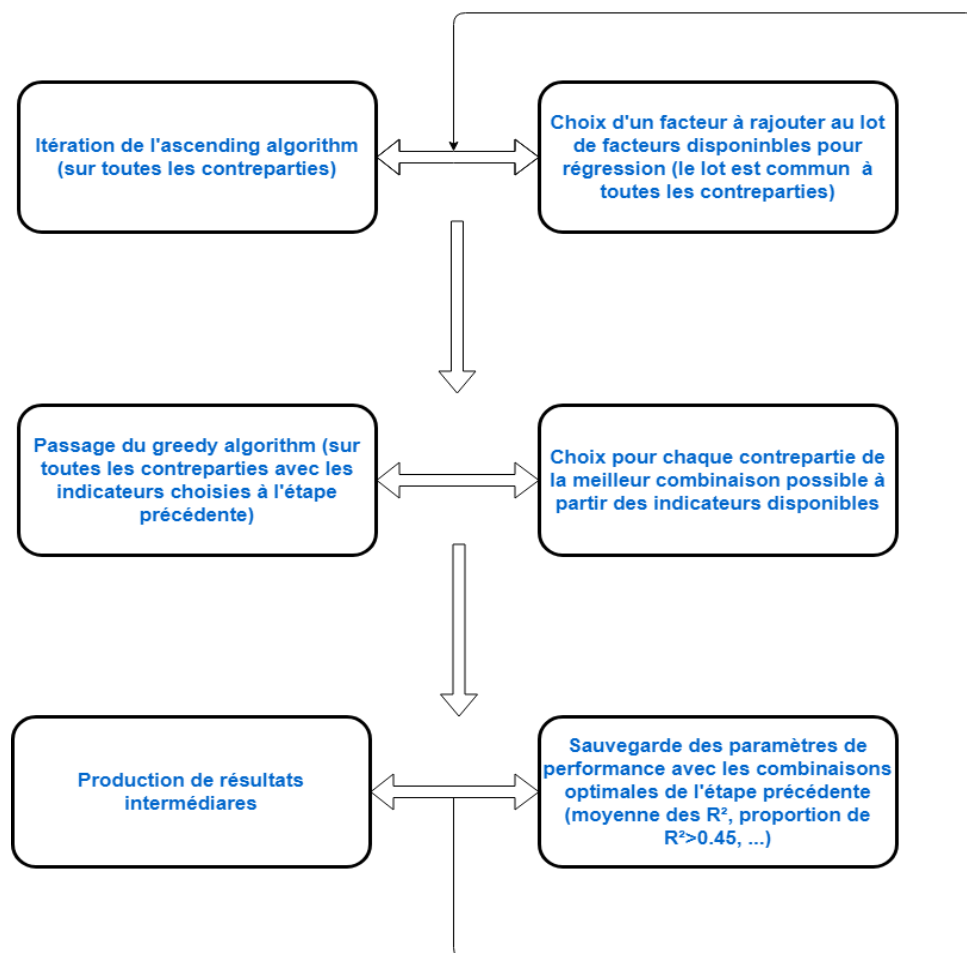


Figure 18: Algorithme de sélection des indicateurs et de calcul des coefficients

³⁰ NB : L'algorithme ascendant est un algorithme de sélection des indicateurs. Sa seule fonction est de choisir quels indicateurs conserver. Lorsqu'il a choisi quels indicateurs utiliser, il est nécessaire d'utiliser un autre algorithme afin de chercher quel est la meilleure régression que l'on peut obtenir à partir des facteurs sélectionnés

Les paramètres retenus pour l'analyse de l'effet factoriel sont les suivants. Ils sont liés à deux caractéristiques des régressions : le coefficient de détermination R^2 et l'écart-type des erreurs σ_ε .

Il s'agit de :

- La somme des R^2 ;
- La somme des R^2 ajustés ;
- La somme des σ_ε ;
- La proportion des $R^2 > 0.45$;
- La proportion des R^2 ajustés > 0.35 ;
- Le nombre de contreparties analysées.

Une attention importante doit être portée au dernier paramètre. En effet, le nombre de contrepartie analysée, ne correspond pas nécessairement au nombre de contreparties de l'échantillon.

En effet tous les indicateurs macroéconomiques ne sont pas disponibles pour tous les pays.

Par exemple, si le Total des Réserves de l'Allemagne n'est pas disponible et qu'au premier itéré, l'algorithme ascendant choisi le Total des Réserves, aucune contrepartie Allemande ne sera incluse.

À la fin de l'algorithme, on connaît le nombre de contreparties maximal analysées (il est atteint au dernier itéré).

Cela permet de définir une 'moyenne pénalisée' des sommes. Plutôt que de diviser par le nombre de contreparties analysées à l'itéré (ce qui donnerait une vraie moyenne), on divise par le nombre de contreparties maximal analysé.

Résultats :

À ce point, les experts en économie impliqués dans le process de construction du modèle de risque de crédit interne à l'IFI, nous ont recommandé de rajouter des indicateurs lagés (retardés). En effet, certains indicateurs macroéconomiques accusent un délai dans leur fluctuation par rapport aux observables du marchés, notamment les ratings.

La technique que nous avons appliquée pour inclure très simplement les facteurs lagés de 1 an et de 2 ans a été de les considérer comme des indicateurs à part entière et de les faire passer par le même processus de sélection que tous les indicateurs.

Ainsi le panel d'indicateurs passe donc à 60 (20 x 3) indicateurs.

Les indicateurs qui sont ressortis sont les suivants :

- Les Taux de changes lagés de 2 ans
- Les Taux 3 mois, lagés d'1 an
- Le PIB par habitant (\$ US courants) sans lags
- Les Dépenses nationales brutes lagées de 2 ans
- La Croissance du PIB (% annuel)

On remarquera la grande cohérence des résultats, avec aucune redondance en terme économique (Croissance du PIB + Croissance du PIB/Habitant par exemple) mais aussi au sens quantitatif (Les Taux 3 mois, lagés d'1 an + Les Taux 3 mois, lagés de 2 ans).

	Mean RSQ	Mean RSQ Adj	Sum Sigma Eps	%RSQ > 0.45	%RSQ Adj > 0.35	% CPs Involved
1 Indicator	24.13%	15.17%	188.36	10.67%	27.20%	98.54%
2 Indicators	38.61%	30.15%	167.39	27.41%	51.46%	98.54%
3 Indicators	49.40%	39.95%	148.89	41.42%	66.74%	100.00%
4 Indicators	58.35%	48.58%	135.37	55.44%	75.10%	100.00%
5 Indicators	66.07%	56.32%	121.02	64.23%	83.26%	100.00%

Tableau 14: Evolution de la qualité des régression en fonction du nombre d'indicateurs disponible

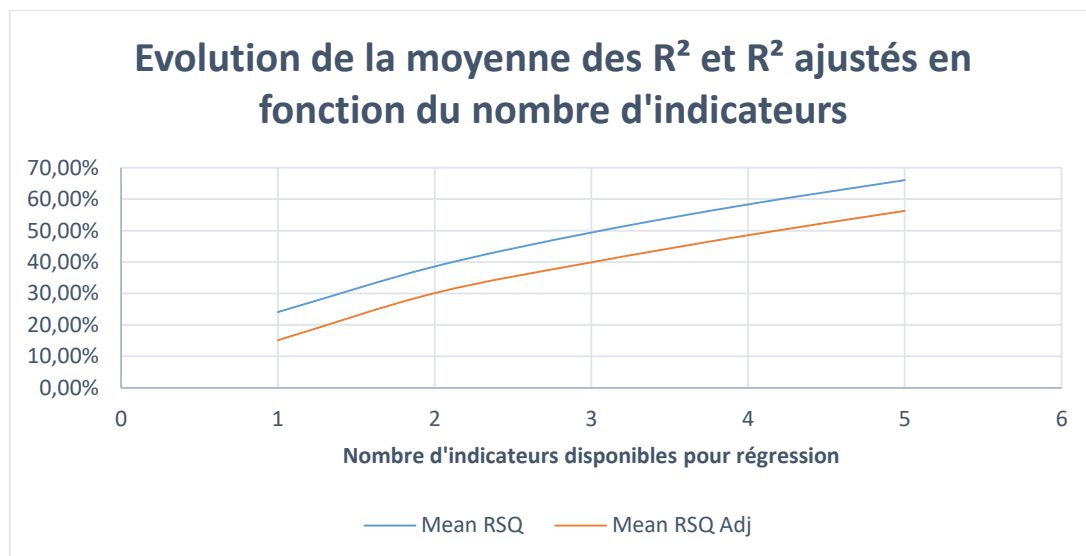


Figure 19: Evolution des R^2 et R^2 ajustés en fonction du nombre d'indicateurs disponible:

4. La gestion des chocs et Stress scénarios

Nous souhaitons mesurer la robustesse des contreparties à des événements de chocs (ou de stress) ; mais aussi celle de l'IFI qui doit pouvoir se couvrir face à de tels événements.

Dans les stress scénarios, outre le fait de choquer les notes ou d'autres paramètres, on va vouloir **fixer a priori des valeurs sur tout ou partie des facteurs**. Par exemple, on souhaitera fixer la variation du PIB d'un pays ou d'un groupe de pays et regarder comment se comporte alors les variations de DD sur ces nouvelles bases.

Dans l'expression (20) :

$$\Delta Z_i(t) = \mu_{Z_i} + \sigma_{Z_i} \left[\sqrt{1 - RSQ_{zi}} \cdot Y(0,1) + \sqrt{RSQ_{zi}} \cdot \sum_{k=1}^m U_{k(j)} \alpha_{ik} \right]$$

La composante idiosyncratique de la variation n'est pas impactée par un choc concernant un ou plusieurs des facteurs, mais la composante factorielle est changée du fait d'avoir fixé les valeurs de facteurs pour un sous ensemble de facteurs et de pays.

Expression d'une distribution normale multivariée conditionnelle

Soit $\vec{\Delta f} = N[\vec{\mu}, \Sigma]$ la distribution multivariée de départ des couples pays-facteurs de dimension (mxp).

On prend (r) couples de 'pays-facteurs' dont nous fixons la valeur des facteurs sur le vecteur $\langle \vec{a} \rangle$ et (q) couples de 'pays-facteurs' laissés libres avec $+q = m \times p$. On peut alors adopter la notation suivante :

$$\langle \vec{\Delta f} \rangle = \begin{pmatrix} \vec{\Delta f}_1 \\ \vec{\Delta f}_2 = \vec{a} \end{pmatrix} \quad \langle \vec{\mu} \rangle = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \quad \text{et} \quad \langle \Sigma \rangle = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$$

Avec $\vec{\Delta f}_1$ et $\vec{\mu}_1$ des vecteurs de dimension q, $\vec{\Delta f}_2$ et $\vec{\mu}_2$ des vecteurs de dimension r et Σ_{11} une sous-matrice carrée (q,q), Σ_{22} une sous-matrice carrée (r,r), Σ_{12} une sous-matrice (q,r), Σ_{21} une sous-matrice (r,q).

Le théorème de Kalman-Bucy relatif aux distributions normales multivariées conditionnelles permet d'établir que la distribution des (q) variables laissées libres $\vec{\Delta f}_{1|a}$ par rapport aux (r) valeurs de $\vec{\Delta f}_2$ fixées sur le vecteur $\langle \vec{a} \rangle$ suit une distribution multivariée $N[\mu_{1|a}, \Sigma_{11|a}]$ avec :

$$\begin{cases} \mu_{1|a} = \mu_1 + \Sigma_{12} \cdot \Sigma_{22}^{-1} \cdot (a - \mu_2) \\ \Sigma_{11|a} = \Sigma_{11} - \Sigma_{12} \cdot \Sigma_{22}^{-1} \cdot \Sigma_{21} \end{cases}$$

Si on souhaite utiliser la matrice de corrélation et pas la matrice de covariance, il faut alors prendre en compte la relation : $\langle \Sigma \rangle = \langle \text{Diag } \sigma \rangle \langle \rho \rangle \langle \text{Diag } \sigma \rangle$ qui conduit à :

$$\langle \Sigma \rangle = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} = \begin{pmatrix} \sigma_{11} \rho_{11} \sigma_{11} & \sigma_{11} \rho_{12} \sigma_{22} \\ \sigma_{22} \rho_{21} \sigma_{11} & \sigma_{22} \rho_{22} \sigma_{22} \end{pmatrix}$$

Avec $\sigma_{11} = \langle \text{Diag } \sigma_1 \rangle$ Avec $\sigma_{22} = \langle \text{Diag } \sigma_2 \rangle$

$$\text{Et } \langle \sigma \rangle = \begin{pmatrix} \sigma_1 \\ \sigma_2 \end{pmatrix}$$

Soit :

$$\left\{ \begin{array}{l} \mu_{1|a} = \mu_1 + \Delta\mu_{1|a} \text{ avec } \Delta\mu_{1|a} = \sigma_{11} \rho_{12} \cdot \rho_{22}^{-1} \cdot \langle \text{Diag } 1/\sigma_2 \rangle \cdot (a - \mu_2) \\ \Sigma_{11|a} = \sigma_{11} \cdot \rho_{11|a} \cdot \sigma_{11} \text{ où } \rho_{11|a} = \rho_{11} - \rho_{12} \cdot \rho_{22}^{-1} \cdot \rho_{21} \end{array} \right.$$

 ρ étant la matrice de corrélations.Application à l'expression de la distance au défaut

À la différence d'une expression mono factorielle, il n'y a plus d'un côté les contreparties dans les pays 'fixés' et de l'autre les contreparties dans les pays laissés libres. Il n'y a qu'un ensemble de (r) couples de 'pays-facteurs' fixés et de (q) couples libres.

Ceci conduit en expression vectorielle à une variation de distance au défaut pour la contrepartie (i) qui répond aux expressions suivantes :

$$\Delta Z_i = \langle \beta_i, \Delta f \rangle + \gamma_i + \varepsilon_i$$

$$\Delta Z_i = \langle \beta_{i1}, \Delta f_1 \rangle + \langle \beta_{i2}, a \rangle + \gamma_i + \varepsilon_i$$

$$\Delta Z_i = \langle \beta_{i1}, \mu_1 \rangle + \langle \beta_{i1}, \Delta\mu_{1|a} \rangle + \langle \beta_{i1}, N[0, \Sigma_{11|a}] \rangle + \langle \beta_{i2}, \mu_2 \rangle + \langle \beta_{i2}, \Delta a \rangle + \gamma_i + \varepsilon_i$$

$$\Delta Z_i = \mu_{Z_i} + \sqrt{1 - RSQ_{zi}} \cdot \sigma_{Z_i} Y(0,1) + \left[\langle \beta_{i1}, \Delta\mu_{1|a} \rangle + \langle \beta_{i2}, \Delta a \rangle + \langle \beta_{i1}, N[0, \Sigma_{11|a}] \rangle \right]$$

En adoptant l'expression normalisée :

$$\vec{\beta}_i = \sqrt{RSQ} \cdot \sigma_{zi} \cdot (\vec{\alpha}_{zi} / \vec{\sigma}_F)$$

En substituant dans le second membre les expressions β_i on obtient :

$$\sqrt{RSQ_{zi}} \cdot \sigma_{Z_i} \left[\langle \alpha_{zi1} / \sigma_{F1}, \Delta\mu_{1|a} \rangle + \langle \alpha_{zi2} / \sigma_{F2}, \Delta a \rangle + \langle \alpha_{zi1} / \sigma_{F1}, N[0, \Sigma_{11|a}] \rangle \right]$$

Soit au final :

$$\Delta Z_i = \mu_{Z_i} + \sigma_{Z_i} \left(\sqrt{1 - RSQ_{zi}} \cdot Y(0,1) + \sqrt{RSQ_{zi}} \cdot \left[\langle \alpha_{zi1} / \sigma_{F1}, \Delta\mu_{1|a} \rangle + \langle \alpha_{zi2} / \sigma_{F2}, \Delta a \rangle + \langle \alpha_{zi1} / \sigma_{F1}, N[0, \Sigma_{11|a}] \rangle \right] \right)$$

$$\text{Si on pose } \Delta\mu_{Z_i} = \sigma_{Z_i} \sqrt{RSQ_{zi}} \cdot \left[\langle \alpha_{zi1} / \sigma_{F1}, \Delta\mu_{1|a} \rangle + \langle \alpha_{zi2} / \sigma_{F2}, \Delta a \rangle \right]$$

$$\Delta Z_i = \mu_{Z_i} + \Delta \mu_{Z_i} + \sigma_{Z_i} \sqrt{1 - RSQ_{zi}} \cdot Y(0,1) + \sigma_{Z_i} \sqrt{RSQ_{zi}} \cdot \left\langle \frac{\alpha_{zi1}}{\sigma_{F1}}, N \left[0, \Sigma_{11|a} \right] \right\rangle$$

Sinon on rappelle que $\Sigma_{11|a} = \sigma_{11} \cdot \rho_{11|a} \cdot \sigma_{11}$ avec $\sigma_{11} = \langle \text{Diag } \sigma_{F1} \rangle$

Soit $\Lambda_{11|a}$ la matrice de Cholesky de $\rho_{11|a}$ et le vecteur gaussien $U_{1|a} = \Lambda_{11|a} Y_q(0,1)$

$$\left\langle \frac{\alpha_{zi1}}{\sigma_{F1}}, N \left[0, \Sigma_{11|a} \right] \right\rangle = \left\langle \frac{\alpha_{zi1}}{\sigma_{F1}}, \sigma_{11} \Lambda_{11|a} Y_q(0,1) \right\rangle = \langle \alpha_{zi1}, U_{1|a} \rangle$$

Il s'agit d'appliquer le produit scalaire avec le vecteur réduit aux q couples de facteurs x pays non fixés soit :

$$\langle \alpha_{zi1}, U_{1|a} \rangle = \sum_{k=1, k \neq r}^{m \times p - r} U_{1|a.k} \alpha_{ik}$$

Si on remarque que par rapport au pays (j) de la contrepartie (i) on a toujours $\alpha_{ik \neq j} = 0$

Ceci conduit à l'expression finale :

$$\Delta Z_i = \mu_{Z_i} + \Delta \mu_{Z_i} + \sigma_{Z_i} \left[\sqrt{1 - RSQ_{zi}} \cdot Y(0,1) + \sqrt{RSQ_{zi}} \cdot \sum_{k=1, k \neq r}^m U_{1|a.k(j)} \alpha_{ik} \right]$$

Cette expression doit faire l'objet d'une renormalisation si on souhaite se caler sur les distributions de matrices selon les principes exposés précédemment.

Cas particulier où le choc de facteurs est opéré sur un pays synthétique qui correspond à la moyenne pondérée de plusieurs pays

Soit $G = \sum \pi_{i*} P_{i*}$ un groupe de n pays ($n \leq p$, le nombre de pays) dont la moyenne d'un des facteurs a été fixée à une unique valeur scalaire a . Les coefficients π_{i*} correspondent aux poids relatifs des pays au sein du groupe.

On introduit dans la distribution des $m \times p$ couples facteurs-pays un nouveau couple Δf_G qui va correspondre à la moyenne pondérée des valeurs du facteur retenu pour les n pays du groupe³¹. Ce nouveau couple facteur-pays synthétique vient en substitution à l'un des couples du facteur considéré au sein des n pays du groupe : C_{pivot}

³¹ Ceci impose de reconstituer la série des valeurs de ce facteur synthétique comme la somme pondérée par les π_{i*} des valeurs du facteur dans les n pays du groupe. Il faudra recalculer les paramètres mu, sigma et la matrice de corrélation associée en conséquence.

Il ne reste plus qu'à appliquer à cette nouvelle distribution multivariée l'expression de la forme conditionnelle de Kalman Buci en situation d'une variable fixée.

Dans l'expression ci-dessous, le couple Δf_G correspondra à Δf_2 et les couples résiduels C_i vont former le vecteur Δf_1 .

$$\langle \Delta f \rangle = \begin{pmatrix} \Delta f_1 = \Delta f_{C_i} \\ \Delta f_2 = a \end{pmatrix} \quad \langle \mu \rangle = \begin{pmatrix} \mu_{C_i} \\ \mu_G \end{pmatrix} \text{ et } \langle \Sigma \rangle = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$$

Dans ce cas, on va obtenir avec la moyenne de G fixée à (a) la distribution multivariée conditionnelle des C_i non fixés :

$N[\langle \mu_{1|a} \rangle, \langle \Sigma_{11|a} \rangle]$ avec les expressions habituelles :

$$\left\{ \begin{array}{l} \mu_{1|a} = \mu_1 + \Delta \mu_{1|a} \text{ avec } \Delta \mu_{1|a} = \sigma_{11} \rho_{12} \cdot \rho_{22}^{-1} \cdot \left(I \frac{1}{\sigma_2} \right) \cdot (a - \mu_2) \\ \Sigma_{11|a} = \sigma_{11} \cdot \rho_{11|a} \cdot \sigma_{11} \text{ où } \rho_{11|a} = \rho_{11} - \rho_{12} \cdot \rho_{22}^{-1} \cdot \rho_{21} \end{array} \right.$$

On remarquera que dans les C_i figure le sous ensemble des n-1 couples C_{i*} du facteur considéré au sein du groupe pays G.

Comme on ne souhaite pas garder Δf_G comme facteur, mais réintroduire la valeur du facteur du pays P_{pivot} on aura pour ce dernier une expression qui dépend des n-1 valeurs des facteurs Δf_{i*} du sous ensemble des C_{i*} :

$$\Delta f_{pivot} = [a - \sum \pi_{i*} \Delta f_{i*}] / \pi_{pivot}$$

En l'état, la manipulation des Δf_{i*} et $\Delta f_{i pivot}$ n'est pas pratique car nous utilisons dans nos calculs de variations des distances au défaut ΔZ_i des aléas centrés normés corrélés issus d'une matrice de Cholesky : $U_{1|a} = \Lambda_{11|a} Y_q(0,1)$

$$\Delta Z_i = \mu_{Z_i} + \Delta \mu_{Z_i} + \sigma_{Z_i} \left[\sqrt{1 - RSQ_{Zi}} \cdot Y(0,1) + \sqrt{RSQ_{Zi}} \cdot \sum_{k=1, k \neq r}^m U_{1|a.k(j)} \alpha_{ik} \right]$$

Il est donc nécessaire d'aller substituer dans le vecteur $U_{1|a}$ la valeur $U_{1|a(pivot)}$ normée centrée issue de Δf_{pivot} selon le processus suivant :

- Tout d'abord prendre en compte que les drifts des n-1 couples C_{i*} dans le Groupe pays sont impactés du fait du choc soit $\mu_{f_{i*}} = \mu_{f_{0i*}} + \Delta \mu_{1|a_{f_{i*}}}$
- Recalculer les valeurs des $\Delta f_{i*} = \mu_{f_{i*}} + \sigma_{f_{i*}} U_{1|a.f_{i*}}$ et de $\Delta f_{pivot} = [a - \sum \pi_{i*} \Delta f_{i*}] / \pi_{pivot}$



- Puis enfin calculer $U_{1|a}(\text{pivot}) = [\Delta f_{\text{pivot}} - \mu_{f_{\text{pivot}}}] / \sigma_{f_{\text{pivot}}}$

Restent à préciser les expressions à retenir pour $\mu_{f_{\text{pivot}}}$ et $\sigma_{f_{\text{pivot}}}$:

$$\left[\begin{array}{l} \mu_{f_{\text{pivot}}} = [a - \sum \pi_{i*} \mu_{i*1|a}] / \pi_{\text{pivot}} \\ \mu_{f_{\text{pivot}}} = [a - \sum \pi_{i*} \mu_{i*} - \sum \pi_{i*} \Delta \mu_{i*1j|a}] / \pi_{\text{pivot}} \end{array} \right.$$

Pour l'écart type $\sigma_{f_{\text{pivot}}}$ on considérera qu'il demeure inchangé par rapport à sa valeur calibrée sur la série historique du facteur concerné.

IV. Exemples de résultats

Variations de la VaR en fonction de la notation :

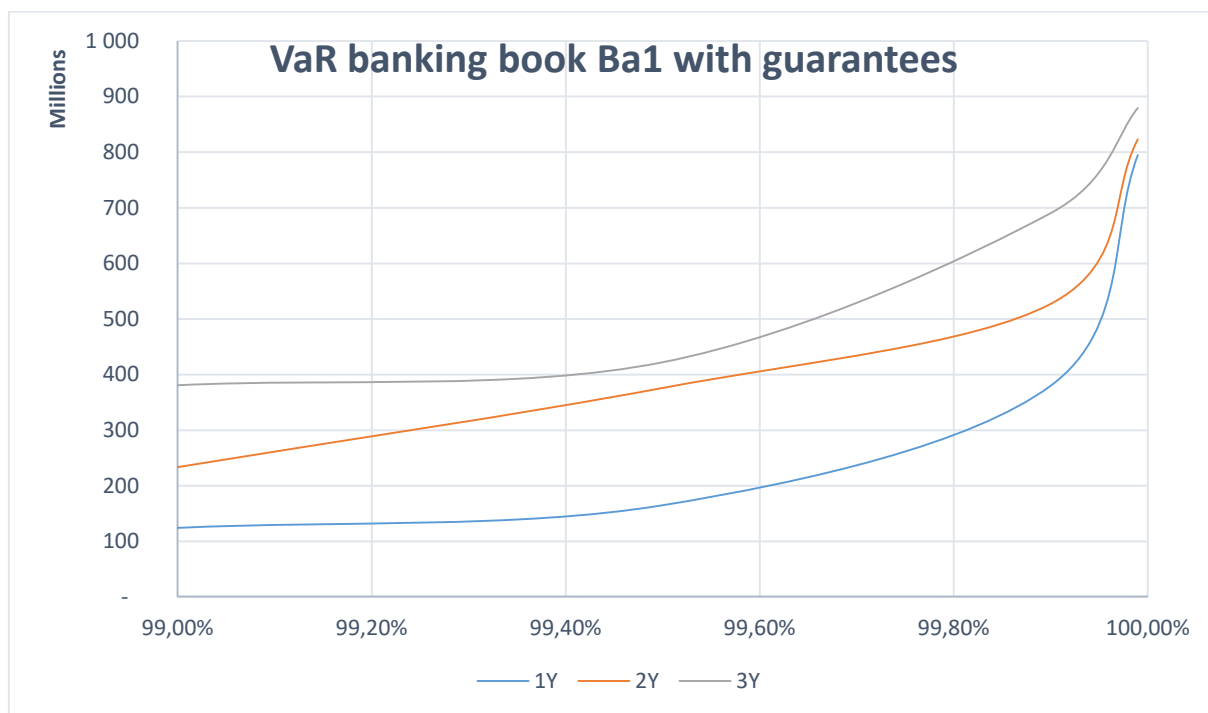


Figure 20: Evolution de la VaR 1Y, 2Y et 3Y avec garanties pour tiers notées Ba1 (Moody's)

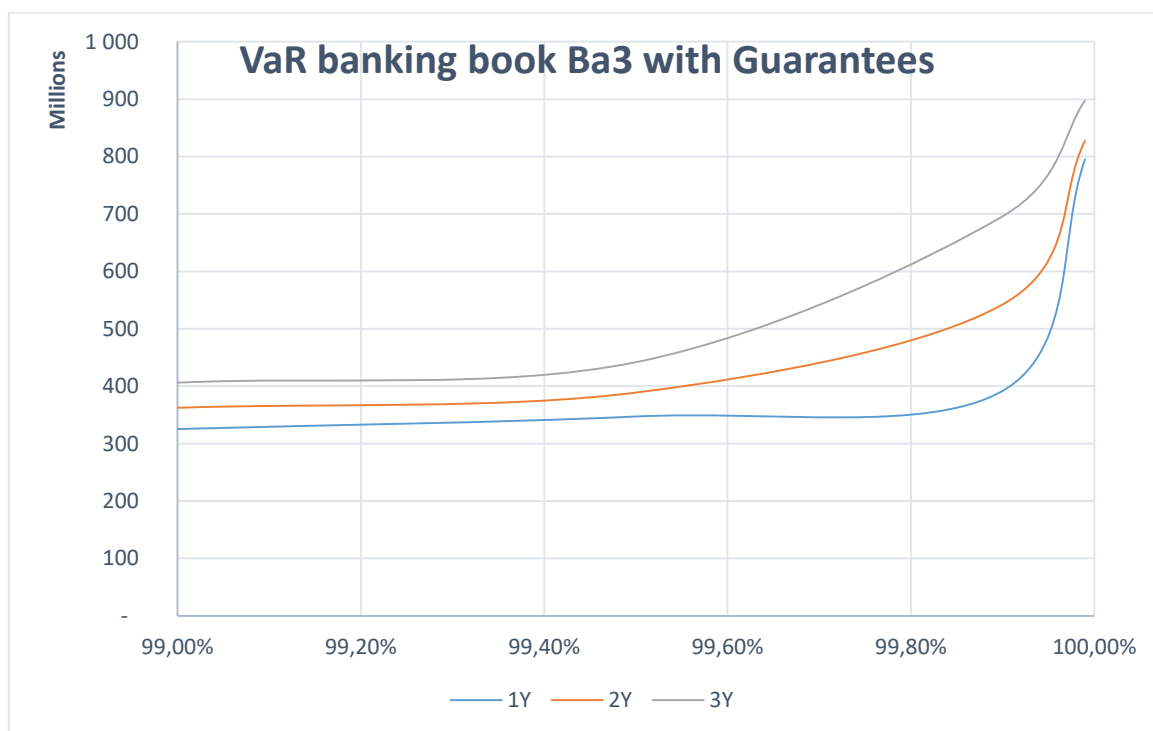


Figure 21: Evolution de la VaR 1Y, 2Y et 3Y avec garanties pour tiers notées Ba3 (Moody's)

Les entreprises notées Ba1 étant relativement plus solvable que celles notées Ba3, on remarque une différence entre les deux sets de graphes. On constate, surtout en horizon 1 an que les pertes sont plus élevées dans le graphe propre aux entreprises Ba3.

Variation de la VaR après la considération des garanties

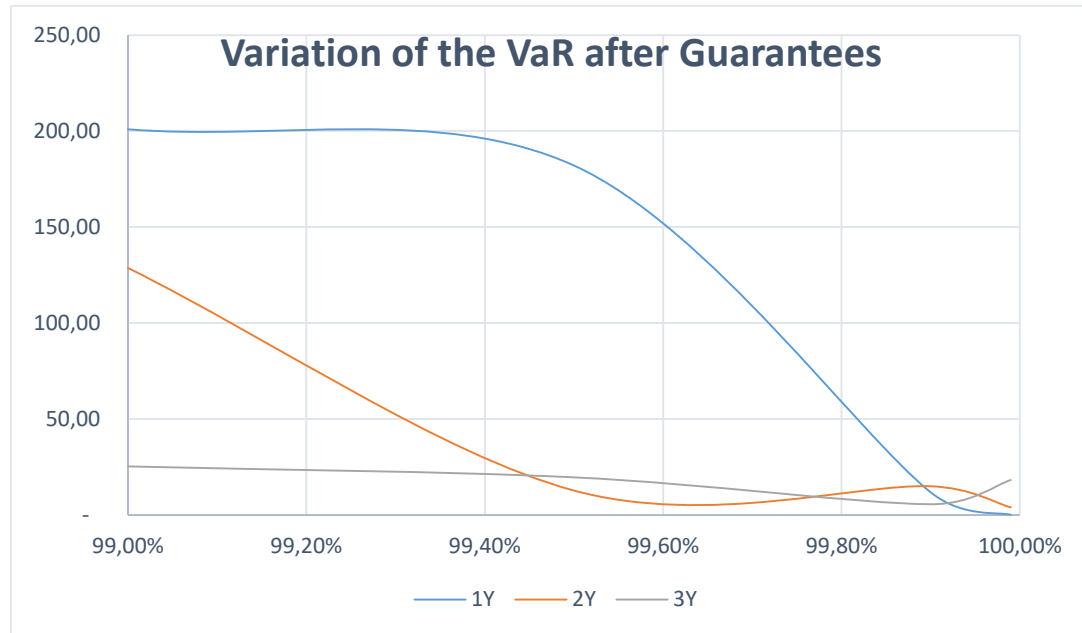


Figure 22: Variation de la VaR après la considération des garanties

On constate facilement comment l'inclusion des garanties (y compris bien sûr des sociétés mères) réduit le nombre de défauts et naturellement la VaR.

Stress sur les LGDs

Nous allons analyser l'effet des chocs de LGD sur la VaR 99.9% avec des horizons de 1 à 5 ans. La LGD va être diminuée de 25% puis de 33%.

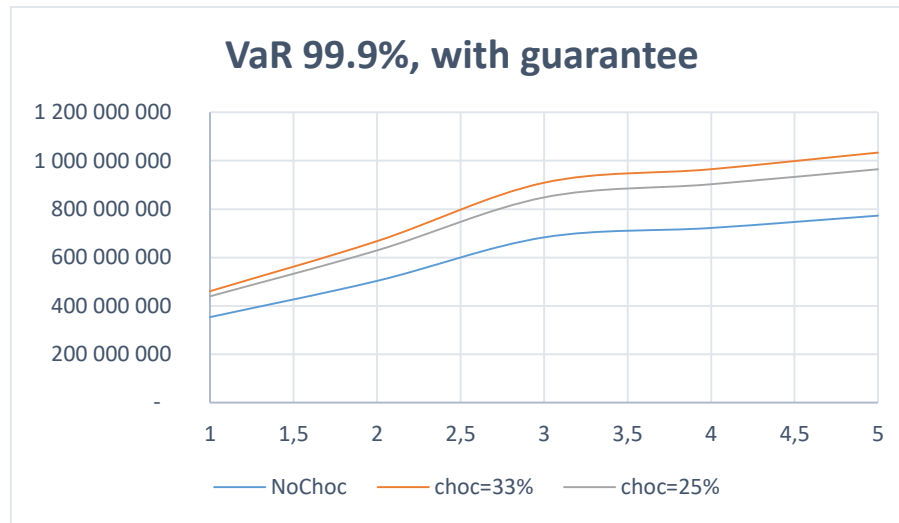


Figure 23: Effet d'un choc de LGD sur la VaR

Naturellement lorsque la LGD augmente, la VaR augmente avec elle. Ce résultat se constate parfaitement à travers le graphe précédent.

Stress sur la croissance du PIB

Nous appliquons un choc - positif - sur la croissance du PIB dans un cas particulier monofactoriel. Nous augmentons de 100% le PIB dans des zones géographiques spécifiques du portefeuille de contreparties.

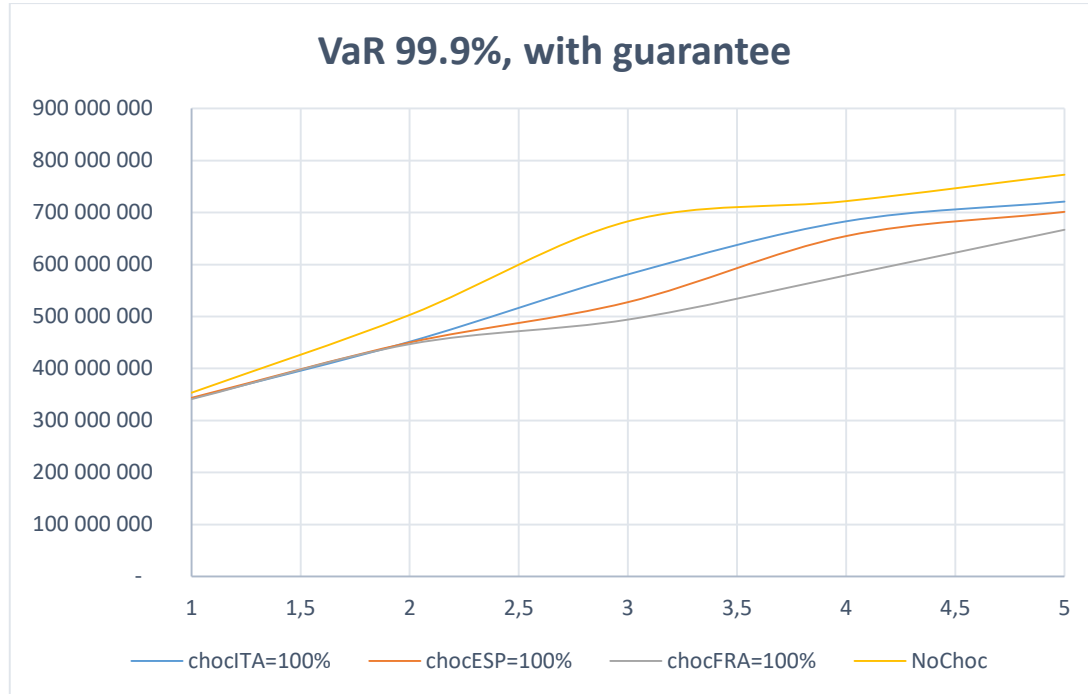


Figure 24: Effet d'un choc positif sur la croissance du PIB

Ainsi, que l'on pouvait si attendre, le fait d'augmenter la valeur de la croissance du PIB, implique une diminution systématique de la VaR simulée.

Evolutions des ratings des CPs

Enfin, nous pouvons également analyser l'évolution des ratings des contreparties de l'IFI sur les horizons 1 an à 3 ans

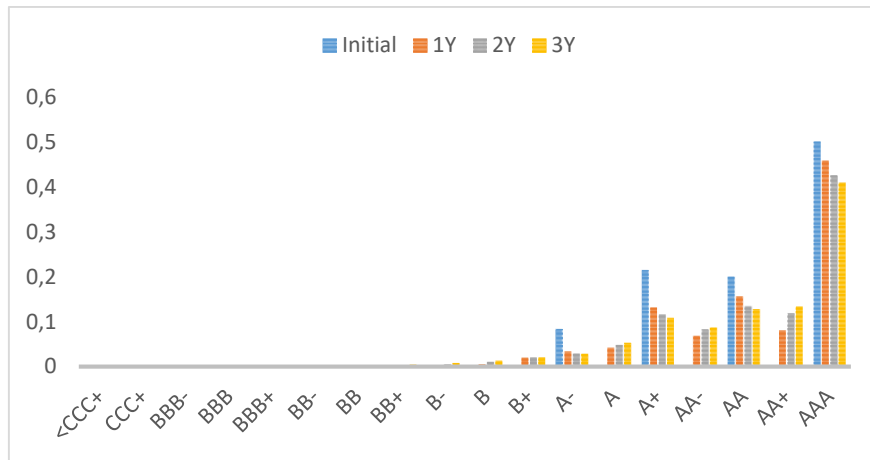


Figure 25: Evolution des ratings des Cps de l'IFI

V. Challenges et perspectives de perfectionnement du modèle

Même si le modèle que nous avons élaboré est tout à fait fonctionnel et donne des résultats tout à fait cohérents et appréciables, le modèle peut toujours être amélioré.

Deux problèmes se posent naturellement, dus aux hypothèses qui sont faites dans le modèle.

- Le premier concerne le passage des ratings aux distances au défaut. Si la fonction affine permettant de passer d'une table à l'autre venait à changer, ce serait naturellement toutes les calibrations déjà effectuées qui deviendraient caduques. Il est donc nécessaire de réfléchir à une solution réelle, basée sur un résultat mathématique, économique ou financier afin de fixer cette fonction définitivement.
- Dans cette même optique, il est nécessaire de fixer la fonction permettant d'affecter le coefficient de corrélation à l'aléa souverain. Il a été évoqué d'utiliser les spreads émetteurs pour cela. La piste reste encore à explorer.

Une autre limite du modèle qui est très peu envisageable de surmonter, sont les hypothèses de normalité. Par exemple, pour les facteurs, nous avons très bien vu, que la NIG était, - à défaut d'être parfaite - plus adaptée que la loi gaussienne. Mais pour des questions de coût de calcul mais aussi de difficultés mathématiques, il paraît peu envisageable de les utiliser.

Conclusion

Au terme de ce stage, nous avons pu concevoir un modèle de risque de crédit sensible aux caractéristiques propres d'une IFI. Les simulations, notamment de VAR, CVAR et des nombre défauts donnent des résultats tout à fait cohérents.

La suite des travaux consistera d'abord en l'implémentation des chocs (factoriels, de LGDs et de PDs). Ensuite, il s'agira naturellement de back-tester de manière beaucoup plus approfondis le programmes ainsi que le modèle. Enfin, les travaux seront clôturés par l'industrialisation des programmes développés pour la calibration du modèle et ceux pour les simulations. Il faudra avec une certaine fréquence (annuelle peut-être) refaire les calibrations des paramètres de régressions. De plus, avec une fréquence moins importante (peut-être de 3 ans), relancer les programmes sélectionnant les indicateurs.

Ce stage a été une très belle et riche expérience. Il a été l'opportunité parfaite pour mettre en application et approfondir des connaissances en probabilités, statistiques et risque de crédit. Il a aussi permis d'obtenir et d'asseoir des connaissances en économies et en finances que nous ignorions jusque-là. Surtout, ce stage a été l'occasion d'avoir une expérience dans le monde de la recherche industrielle.

Comme une ouverture de perspective, il pourrait être intéressant de s'interroger sur la manière d'inclure dans un pareil modèle multifactoriel un autre terme stochastique relié à des données de marchés (similaire au modèle de Merton par exemple avec la valeur de la compagnie) afin que le modèle puisse être adaptable à des Banques et Entreprises avec des contreparties et donc des expositions au risque différentes de celles de notre IFI.

Annexes

A. Régressions linéaires^{viiiix}

Elle est avec l'ANOVA³² ou l'ANCOVA³³ un type de modèle linéaire. C'est-à-dire, un modèle qui peut s'écrire sous la forme :

$$Y = \sum_{j=1}^k \beta_j X^j + E$$

- Y est une variable aléatoire que l'on observe et que l'on souhaite expliquer dans notre cas, mais prédire dans d'autres. On suppose que la variance de Y est constante (hypothèse d'homoscédasticité).
- La famille des $\{X_i, \dots, X_k\}$ est l'ensemble des variables explicatives. Ils sont naturellement eux aussi connus.
- Les β_j $j \in \{1, \dots, k\}$ sont les paramètres du modèle, non observés et donc à estimer par des techniques statistiques appropriées (en fonction des quantités que l'on voudra ou maximiser ou minimiser en général).
- E est le terme d'erreur dans le modèle ; c'est une variable aléatoire réelle non observée pour laquelle on pose les hypothèses suivantes :

$$E(E) = 0 ; Var(E) = \sigma^2 > 0 \text{ où } \sigma^2 \text{ est un paramètre inconnu, à estimer.}$$

Les hypothèses posées sur E impliquent les caractéristiques suivantes sur Y :

$$\begin{cases} E(\hat{Y}) = \sum_{j=1}^k \beta_j X^j \\ Var(\hat{Y}) = \sigma^2 \end{cases} \quad 34$$

Soient :

- y le vecteur de R^n composé des valeurs y_1, \dots, y_n ,
- X la matrice (n, k) de rang k , contenant les valeurs observées des k variables explicatives, disposées en colonnes,
- β le vecteur de R^k contenant les k paramètres du modèle,
- E le vecteur de R^n des erreurs du modèle

Nous obtenons $y = X\beta + E$

³² Analyse de Variance

³³ Analyse de Covariance

³⁴ Les chapeaux désignent les estimés

Pour la régression linéaire multiple avec m facteurs on écrit $y_i = \alpha + \beta_1 x_{i1} + \dots + \beta_m x_{im} + N(0, \sigma_\varepsilon), \forall i \in \{1, \dots, n\}$ ³⁵. L'écriture matricielle correspondante est :

$$\begin{pmatrix} y_1 \\ \vdots \\ y_i \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & \dots & x_{1m} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{i1} & \dots & x_{im} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \dots & x_{nm} \end{pmatrix} \begin{pmatrix} \beta_0^{36} \\ \beta_1 \\ \vdots \\ \beta_i \\ \vdots \\ \beta_m \end{pmatrix} + \begin{pmatrix} E_1 \\ \vdots \\ E_i \\ \vdots \\ E_m \end{pmatrix}$$

Pour trouver la valeur des coefficients β nous utilisons un estimateur des moindres carrés ordinaires.

On note :

$$S = E^t E = \sum_i E_i^2$$

On souhaite résoudre :

$$\frac{\partial S}{\partial a} = 0^{37}$$

$$\begin{aligned} S &= E^t E = (Y - X\beta)^t (Y - X\beta) \\ &= Y^t Y - 2\beta^t X^t Y + \beta^t X^t X \beta \\ \Rightarrow \frac{\partial S}{\partial \beta} &= -2(X^t X)^{-1} + 2(X^t X)\beta = 0 \\ \Rightarrow \hat{\beta} &= (X^t X)^{-1} X^t Y \quad 38 \end{aligned}$$

Afin de mesurer la performance d'un modèle de régression linéaire, il est nécessaire de procéder à une décomposition de la variance.

On admet que celle-ci se décompose comme suit :

$$Var(Y) = Var(\hat{Y}) + Var(E)$$

Remarquons ensuite les relations suivantes :

$$\begin{aligned} SCT &= \sum_{i=1}^n (Y_i - \bar{Y})^2 = (n-1)Var(Y) \\ SCE &= \sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2 = nVar(\hat{Y}) \\ SCR &= \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 = (n-m-1)Var(E)^{39} \end{aligned}$$

On a évidemment avec la décomposition de la variance

$$SCT = SCE + SCR$$

³⁵ On remarquera également qu'il suffit de rajouter une rangée de 1 à la matrice X pour passer d'une écriture à l'autre

³⁶ On note $\alpha = \beta_0$ pour simplifier les notations

³⁷ $E^t E$ est nécessairement convexe

³⁸ Remarquons, que cette expression ne permet pas d'avoir des coefficients β égaux à 0

³⁹ SCT (Somme des Carrés Totaux), SCE (Somme des Carrés Estimés) et SCR (Somme des Carrés Résiduels)

On définit à présent le coefficient de détermination

$$R^2 = \frac{SCE}{SCT} = 1 - \frac{SCR}{SCT} = 1 - \frac{n - m - 1}{n - 1} \times \frac{Var(E)}{Var(Y)}$$

Un important biais sur ce coefficient est automatiquement observé. En effet, la présence du coefficient $-m$ au numérateur implique une augmentation mécanique du coefficient R^2 lorsque le nombre de facteur augmente.

Une méthode classique ayant pour but de contourner ce problème est l'utilisation du $R^2_{ajusté}$.

$$R^2_{ajusté} = 1 - \frac{\left(\frac{SCR}{n - m - 1}\right)}{\left(\frac{SCT}{n - 1}\right)} = 1 - \frac{Var(E)}{Var(Y)}$$

$$R^2_{ajusté} = \frac{R^2(n - 1) - m}{(n - m - 1)}$$

On remarquera cependant que même ce coefficient présente des limites très significatives.

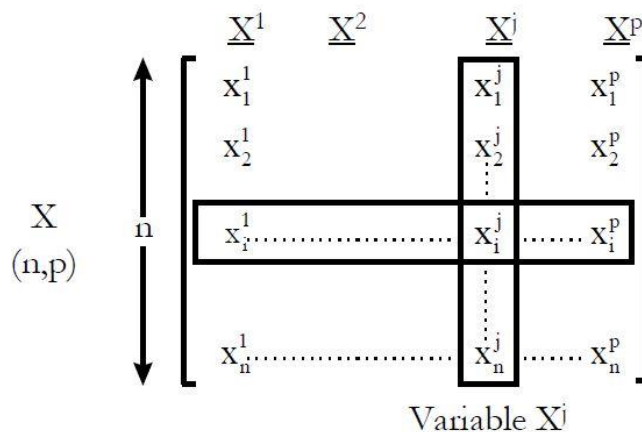
En effet, ce coefficient résulte plus de la volonté d'annuler l'effet du poids $\frac{n-m-1}{n-1}$ que d'un raisonnement mathématique réel. Une des conséquences dommageable est que, alors que ce coefficient est supposé pénaliser l'ajout de facteurs, il n'est pas équivalent au R^2 lorsque le nombre de facteurs vaut 1.

B. PCA

L'Analyse en Composante Principale⁴⁰ permet l'observation des relations entre un ensemble de variables mais aussi la détermination des combinaisons linéaires de variables les plus représentatives de l'échantillon. Elle permet ainsi de procéder entre autres à la réduction de dimensionnalité d'un problème.

Admettons que nous disposions d'un échantillon de n individus observés sur p variables (numériques).

Ces données sont stockées dans une matrice $M \in R^{n \times p}$



Nous cherchons à obtenir une représentation de nos n individus dans le sous-espace $F^k \subset R^p$. Nous cherchons donc k nouvelles variables, appelées composantes principales, qui sont combinaison linéaire des p variables précédentes. On cherche néanmoins à perdre le moins d'information possible. En pratique, on voudra conserver un pourcentage de variance expliquée suffisamment élevé.

Le processus est le suivant.

La matrice M est normalisée⁴¹ pour devenir une matrice M^{norm} de telle sorte que :

$$M_{ij}^{norm} = \frac{M_{ij} - X^i}{\sigma(X^i)}$$

On déduit la matrice de corrélation C :

$$C = \frac{1}{n} M^{norm} (M^{norm})^t$$

Puisque C est symétrique et réelle, elle peut être diagonalisée.

$$\exists P \text{ inversible}, D \text{ diagonale}, C = P^t D P$$

Nous cherchons donc le vecteur u tel que la projection du nuage sur u ait une variance maximale.

⁴⁰ Ou Transformée de Karhunen Loeve

⁴¹ Il est possible de ne pas normaliser la matrice, mais cela implique la possibilité que des variables ayant un drift ou une volatilité plus importante porte tout l'effet de l'ACP sans qu'elle ne soit réellement importante.



La projection de l'échantillon des X sur u s'écrit $\pi_u(M^{norm}) = M^{norm}.u$

La variance empirique de $\pi_u(M^{norm})$ vaut donc :

$$\begin{aligned}\pi_u(M^{norm})^t \cdot \frac{1}{n} \cdot \pi_u(M^{norm}) &= u^t \cdot M^t \cdot \frac{1}{n} \cdot M u \\ &= u^t C u \\ &= (Pu)^t D (Pu) \\ &= v^t D v\end{aligned}$$

Avec $v = Pu$.

Puisque les valeurs propres sont rangées dans l'ordre décroissant.

Le vecteur maximisant l'expression de la variance est le vecteur propre de C associé à λ_1 .

Le processus est réitéré sur les autres axes de projection souhaités en respectant l'orthogonalité.

C. K-means Clustering

Le clustering⁴² est une méthode d'apprentissage non supervisé consistant à la classification en k groupes de n individus $\{X_1, \dots, X_n\}$ observés bien sûr, $k < n$. On souhaite que ces k groupes aient bien sûr des caractéristiques le plus similaire possible et que ceux qui ne seront pas dans la même partition soient le plus différent possible.

Le processus est généralement le suivant :

- On choisit tout d'abord le nombre de groupes que l'on souhaite former ;
- On partitionne d'abord (souvent aléatoirement) les données ;
- En fonction du critère que l'on souhaite optimiser, on déplace les objets d'une classe à une autre ;
- Lorsque la convergence est atteinte on renvoie le clustering final.

Pour le k-means clustering, chaque individu est placé dans le cluster de centre de gravité μ le plus proche.

Soit S l'ensemble des clusters = $\{S_1, S_2, \dots, S_k\}$, on cherche à minimiser la distance entre les points au sein du même cluster.

On cherche donc le set S^* tel que :

$$S^* = \underset{S}{\operatorname{argmin}} \sum_{i=1}^k \sum_{X_j \in S_i} \|X_j - \mu_i\|^2 = \underset{S}{\operatorname{argmin}} \sum_{i=1}^k |S_i| \operatorname{Var} S_i$$

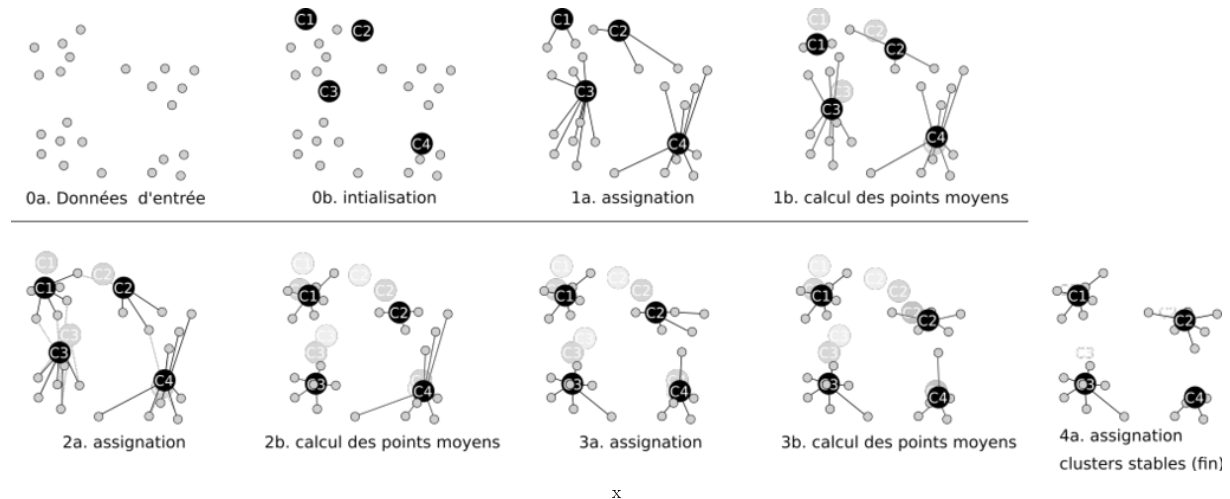


Figure 26: Processus de l'algorithme k-means Clustering

⁴² Partitionnement

D. Recherche Documentaire sur le process de Notation de S&P

Dans l'optique de trouver des indicateurs cohérents économiquement pour expliquer des variations de distance au défaut, une importante recherche bibliographique a été menée. Elle s'est principalement tournée vers la documentation – non payante – de Standard & Poors.

De cette recherche documentaire nous avons tiré 3 conclusions:

- La manière dont les notes sont calculées inclue beaucoup de préoccupations si bien qu'il paraît très peu probable de parvenir à des R^2 élevés en ne considérant qu'un seul indicateur.
- Nos meilleurs indicateurs ne sont pas incohérents par rapport à ce qui se fait.
- Enfin, il pourrait être utile de s'intéresser à la série d'indicateurs Doing Business de la banque mondiale mais aussi aux indicateurs de marché, notamment les cours de change⁴³.

Sovereign Issuer Credit Rating Framework

Five Key Areas To Determine A Sovereign's Creditworthiness

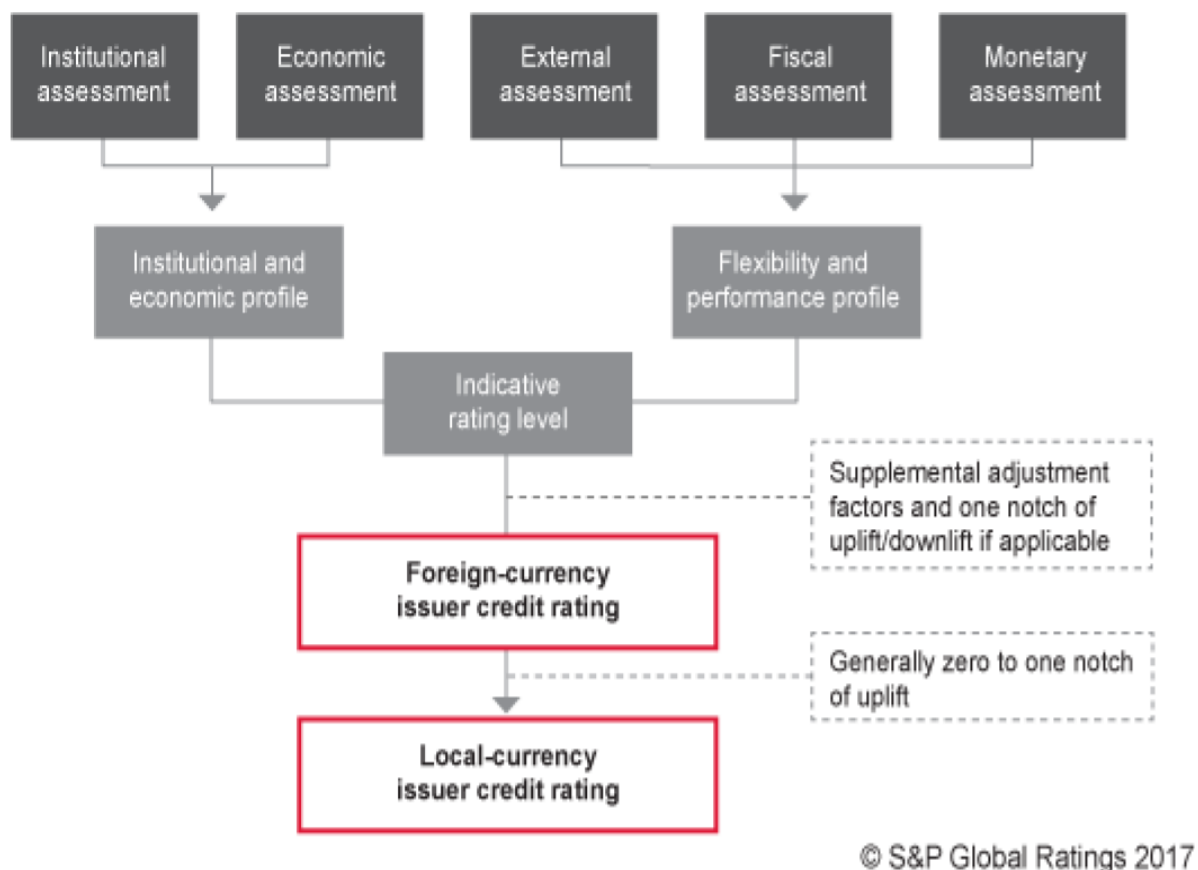


Figure 27: Trame de rating des souverains de S&P

⁴³ En effet finalement inclus dans le panel de 20 indicateurs

Evaluation Economique

<<Les principaux moteurs de notre évaluation économique d'un souverain sont: son niveau de revenu, sa prévisions de croissance, sa diversité et volatilité de l'économie. >>

<<Le PIB par habitant est notre mesure de revenu la plus importante >>.

<<La mesure clé de la croissance économique est la croissance tendancielle du PIB réel par habitant.

<<<Notre évaluation économique utilise l'estimation du PIB par habitant pour l'année en cours tirée des statistiques nationales, convertie en dollars américains>>>.

Nous arrivons donc à très bien incorporée la partie économie de la notation des souverains grâce à l'indicateur PIB en US\$ Courants.

Evaluation Externe

<<Trois facteurs déterminent l'évaluation externe d'un pays :

- Le statut de sa monnaie dans les transactions internationales ;*
- La liquidité extérieure du pays, qui donne une indication de la capacité de l'économie à générer les devises nécessaires pour remplir ses obligations publiques et privées envers les non-résidents ;*
- La position extérieure du pays, qui montre les actifs et les passifs des résidents (...) par rapport au reste du monde. >>*

<<<Notre principale mesure de la liquidité extérieure d'un pays est le rapport entre les besoins extérieurs bruts de financement et la somme des recettes du compte courant (RCA) plus les réserves officielles de change utilisables. >>

<<< L'endettement extérieur du pays est le ratio de la dette extérieure nette étroite par rapport à la RCA.>>

Nous disposons également d'indicateurs à même de capturer cet effet. En effet, le total des réserves, les balances de paiements courants mais aussi la capacité(+)/besoin(-) de financement correspondent bien au descriptif.

Evaluation Fiscale

Évaluation financière

<<L'évaluation fiscale reflète la soutenabilité des balances fiscales et du poids de la dette d'un souverain. Il considère : (...)

- Les performances fiscales et la flexibilité*
- Le poids de la dette >>.*

<<<Notre principale mesure de la performance d'un gouvernement est la variation de l'encours net de la dette publique exprimée en pourcentage du PIB.

<<(<(...) Les conditions suivantes améliorent généralement l'évaluation de la performance d'un souverain et la flexibilité (...):

Les administrations publiques disposent d'importants actifs liquides (généralement plus de 25 % du PIB) pour atténuer les effets des cycles économiques. (...)>>

<<<Une évaluation du poids de la dette d'un souverain reflète son niveau d'endettement potentiel. Les facteurs qui sous-tendent l'évaluation sont les suivants :

- *Dette par rapport au PIB,*
- *Le coût d'intérêt de la dette par rapport aux recettes des administrations publiques,*
- *Structure de la dette et accès au financement >>.*

Ici aussi, nos indicateurs, en l'occurrence les créances sur le gouvernement central ainsi que le total des réserves sont cohérents par rapport aux critères S&P.

Evaluation Monétaire

<<L'évaluation monétaire d'un souverain résulte de l'analyse du :

- *Régime de change, qui permet à un souverain de mener une politique monétaire indépendante et adaptée aux besoins de son économie nationale ; et*
- *La crédibilité de sa politique monétaire, mesurée, entre autres, par les tendances sur un cycle économique ou par l'impact des mécanismes monétaires de marché sur l'économie réelle, qui est fonction de la profondeur et de l'ampleur du système financier local et des marchés des capitaux >>.*

<<<L'évaluation monétaire initiale d'un souverain est obtenue en combinant nos évaluations du :

- *Régime de change (pondéré à 40 %), et*
- *De la crédibilité de la politique monétaire (pondérée à 60 %) >>.*

<<<Un souverain pourrait recevoir une note de devise étrangère inférieure à la note de référence en raison d'une liquidité externe extrêmement faible>>>

Cet aspect semble quelque peu emporté par l'inflation. Cependant avant l'intégration des taux de changes parmi les indicateurs disponibles, cette dimension n'était pas adéquatement couverte.

Evaluation Institutionnelle

<< L'évaluation institutionnelle initiale d'un souverain combine les éléments suivants : l'efficacité, la stabilité et la prévisibilité de ses institutions politiques et décisionnelles ; et la transparence et la responsabilité de ses institutions, données et processus, et la couverture et la fiabilité de l'information statistique.>>

Actuellement, cette dimension est la seule pour laquelle aucun de nos indicateurs ne peut participer à la capture de l'information.

Au-delà de ces piliers, la recherche documentaire nous pousse à nous pencher sur l'acquisition d'autres données.

Données de marchés

[À propos de l'évaluation monétaire]

<<L'évaluation monétaire tient compte de notre appréciation sur la capacité de l'autorité monétaire à remplir sa mission tout en soutenant une économie équilibrée et en atténuant tout choc économique majeur ou tout choc financier. Nous dérivons l'évaluation monétaire en analysant :

- *Le régime de change,*
- *(...) ; et*
- *La crédibilité de la politique monétaire telle qu'elle est mesurée (...) par (...) les mécanismes monétaires axés sur le marché de l'économie réelle, qui est largement fonction de la profondeur et de l'ampleur du système financier et des marchés financiers d'un pays. >>*

Nous percevons bien ici, l'importance des données de marché.

Indicateurs "Doing Business"

[À propos de l'évaluation externe]

<<< Nous pourrions généralement appliquer un ajustement négatif supplémentaire si nous avons des raisons de croire ce qui suit :

- *(...)*
- *(...)*
- *D'autres facteurs externes ou l'évolution de la situation dans le pays peuvent nuire à la réputation du pays en tant que centre international financier stable, ce qui pourrait diminuer son attrait et entraîner une détérioration de l'économie locale, de l'emploi ou des revenus du gouvernement >>.*

Nous avons ici un exemple parmi ceux où on pourrait considérer l'utilisation d'indicateurs donnant une idée du climat des affaires.

E. Distribution NIG

La NIG est une loi à densité continue. Elle est particulièrement appréciée pour ces paramètres supplémentaires qui lui permettent d'avoir une queue de distribution épaisse (kurtosis) ainsi qu'une asymétrie (skewness) d'un côté ou de l'autre.

Elle possède donc 4 paramètres :

- μ : Le paramètre de localisation ;
- δ : Le paramètre d'écart (typiquement σ pour une loi normale) ;
- α : Le paramètre définissant l'épaisseur de la queue ;
- β : Le paramètre d'asymétrie.

La loi de densité est la suivante :

$$\frac{\alpha \delta K_1 \left(\alpha \sqrt{\delta^2 + (x - \mu)^2} \right)}{\pi \sqrt{\delta^2 + (x - \mu)^2}} \exp^{\delta \gamma + \beta(x - \mu)}$$

Avec $\gamma^2 = \alpha^2 - \beta^2$

K_1 est une fonction de Bessel modifiée

$$K_n(x) = \frac{\pi}{2} \times \frac{I_{-n}(x) - I_n(x)}{\sin(n\pi)}$$

$I_n(x)$ est une fonction de Bessel du premier ordre

$$I_n(z) = \frac{1}{2\pi i} \oint e^{\left(\frac{z}{2}\right)^{\frac{(t+1)}{t}}} t^{-n-1} dt$$

Enfin, soit $X \sim NIG(\mu, \delta, \alpha, \beta)$

$$E(X) = \mu + \frac{\delta\beta}{\gamma}$$

$$Var(X) = \frac{\delta\alpha^2}{\gamma^3}$$

$$Skewness = \frac{3\beta}{\alpha\sqrt{\delta\gamma}}$$

$$Kurtosis = \frac{3}{\delta\gamma} \left(1 + \frac{4\beta^2}{\alpha^2} \right)$$

Références

- ⁱ Basel II retail modelling approaches PD Models
- ⁱⁱ Mémoire d'actuaire Ahmed Salim Najit – Université paris Dauphine
- ⁱⁱⁱ Distance-to-Default in Banking: A Bridge Too Far? - Jorge A. Chan-Lau and Amadou N.R. Sy
- ^{iv} A Comparison of Credit Risk Models Risk Theory Enrique Benito, Silviu Glavan & Peter Jacko
- ^v Tomasz Zieliński University of Economics in Katowice - Merton's and KMV models in credit risk management
- ^{vi} Computing the nearest Correlation Matrix – N. Higham 2002 –Manchester Institute for Mathematical Sciences
- ^{vii} Anas Yassine, Abdelmadjid Ibenrissoul, Les déterminants macroéconomiques de la défaillance des emprunteurs: Cas d'une banque marocaine
- ^{viii} C.Chouquet modeles lineaires - Laboratoire de Statistique et Probabilités – Université Paul Sabatier
- ^{ix} Ricco Rakotomalala - Université Lumière Lyon 2
- ^x Par Mquantin — Travail personnel, CC BY-SA 4.0,