

Bâle II

Comment mettre en œuvre un *backtesting* des outils de notation de la banque de détail ?

Un outil de *backtesting* s'assure sur le long terme de la performance des modèles de notations internes et du respect des obligations réglementaires vis-à-vis de Bâle 2. L'investissement n'est pas négligeable mais rentable au vu des enjeux.

LE DÉVELOPPEMENT DE MODÈLES internes de notation, indispensables pour la réforme Bâle II en méthode IRB avancée, ne constitue pas une fin en soi. Une fois ces modèles mis en place, les banques se doivent d'entretenir et de faire vivre leurs outils, non seulement au service d'une bonne gestion opérationnelle et commerciale, mais également parce qu'il s'agit là d'une obligation réglementaire, prévue dans le pilier 2 de la réforme bâloise.

Dès lors, il s'agit de définir clairement des normes ainsi qu'un protocole de *backtesting* : que cherche-t-on à contrôler, comment et à quelle fréquence, par qui et à quel coût ?

LES ENJEUX DU BACKTESTING

Au-delà de la mise en conformité avec les exigences du Comité de Bâle, les enjeux du *backtesting* sont doubles : s'assurer de la qualité des prévisions de défaillance et, le cas échéant, refondre ou faire évoluer les modèles de notation, de façon à prévoir toujours au mieux le défaut. En plus de ces enjeux techniques mais néanmoins ré-

glementaires, le *backtesting* participe à la communication non seulement auprès des directions opérationnelles mais aussi auprès de la Commission bancaire. En effet, il revient à l'entité qui produit le *backtesting* de démontrer aux Directions commerciales que les outils de risque, destinés soit à l'octroi de produits soit à l'évaluation dynamique du comportement des clients, sont bien calibrés, permettant à la banque de bien appréhender les risques qu'elle prend. Quant à la Commission bancaire, elle est en droit d'exiger à tout moment la preuve que les outils de la

banque fonctionnent correctement et font l'objet d'une attention particulière (encadré 1).

DÉFINIR UN PROTOCOLE DE SUIVI DES OUTILS

Afin de répondre correctement aux enjeux du *backtesting*, il convient de définir précisément un protocole dressant la méthodologie du *backtesting*, sa périodicité ainsi que les compétences nécessaires pour le produire au moindre coût.

• Méthodologie

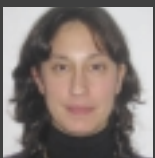
Dans la banque de détail, pour le marché des particuliers et des

1. Le suivi des outils de risque : une obligation réglementaire

« Les banques doivent disposer d'un système solide pour valider l'exactitude et la cohérence des systèmes et procédures de notation et de l'estimation de tous les grands facteurs de risque. Elles doivent démontrer à leur autorité de contrôle que leur processus de validation leur permet d'évaluer, de manière cohérente et significative, les performances de leurs systèmes internes de notation et d'appréciation du risque. Les banques doivent comparer régulièrement les taux de défaut enregistrés avec les estimations PD¹ pour chaque catégorie de notation et être en mesure de prouver que ces taux respectifs res-

tent dans les limites prévues. Dans le cadre de l'approche NI complexe [IRB avancée], elles doivent effectuer cette analyse pour leurs estimations PCD [LGD] et ECD [EAD]. Ces comparaisons doivent s'appuyer sur les données historiques remontant aussi loin que possible. Les méthodes et données utilisées à cet effet doivent être clairement expliquées. L'analyse et les documents doivent être mis à jour au moins une fois l'an. » (Extrait du texte réglementaire Bâle II d'avril 2003 - § 463 et 464).

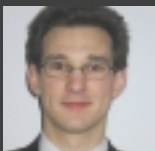
¹ PD : Proba de défaut, LGD : perte en cas de défaut, EAD : encours au moment du défaut.



NATACHA JOURDAN
Risque de crédit



LAURENT MICHEL
Groupe de recherche opérationnelle



ERIC SALOMON
Groupe de recherche opérationnelle

Direction des risques du Groupe
Crédit Agricole SA

Éléments techniques du *backtesting*

La mise en place de procédures normalisées de *backtesting* exige de définir certains indicateurs synthétiques et pertinents qui aident l'analyste à préciser ses diagnostics, via des seuils d'alerte correctement calibrés sur les modèles internes. Nous présentons ci-contre quelques uns des indicateurs couramment utilisés :

La mesure de performance

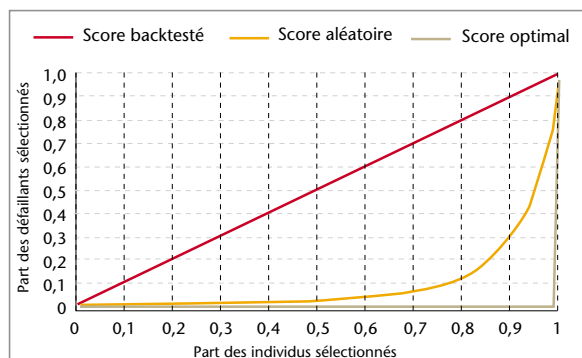
Cette mesure peut être réalisée grâce à la ROC Curve (Receiver Operating Characteristics), couramment utilisée par les agences de rating et qui donne une représentation graphique des performances statistiques d'un score. Elle illustre le pouvoir discriminant d'un outil de notation (graphique).

L'indicateur AR (Accuracy Ratio) vient préciser le diagnostic de performance, en donnant une mesure agrégée. Ce ratio compare la courbe de sélection du modèle testé à celle du modèle idéal re-

présenté par la courbe optimale. Plus précisément, il correspond au rapport de la surface délimitée par le modèle testé et le modèle naïf (tirage au hasard des notes) et celle comprise entre le modèle idéal et ce même modèle aléatoire. Ce ratio, compris entre 0 et 1, permet ainsi de suivre l'évolution des performances d'un outil entre deux périodes de suivi consécutives, ou entre une période de suivi et la période de développement de l'outil.

La mesure de stabilité

Le diagnostic de stabilité est généralement dérivé d'une série d'indicateurs qui précise la déviation de la population scorée par rapport à une population de référence, généralement la population au moment du développement du score ou de sa mise en production. Ils permettent de détecter les changements structurels de la population sur laquelle est appliqué le score, ou la fluctuation de



Exemple de lecture :

Sélectionner 90 % des cas les mieux scorés permet de ne sélectionner que 30 % des défaillants. Autrement dit, 70 % des défaillants figurent parmi les 10 % de scores les plus faibles. Le score correspond à un Accuracy Ratio de 0,82 soit un score de très performant.

contribution d'une variable dans le système de notation.

Un des indicateurs les plus couramment utilisés pour la détection d'instabilité est l'Information ratio, qui s'appuie sur la notion d'entropie statistique.

La mesure de déviation des PD

Cette phase du *backtesting* consiste principalement à confronter les taux de défaut théoriques, prévus au moment du développement de l'outil de notation, et

les taux de défaut réalisés sur la période analysée. Il s'agit de s'assurer que les classes de risque définies restent bien calibrées donc discriminantes, c'est-à-dire ordonnées (elles doivent caractériser une croissance ou décroissance du risque, selon le sens choisi) et deux à deux distinctes. Enfin, on vérifiera, notamment à l'aide de tests statistiques appropriés, que les taux de défaut prévus et constatés, pour chacune des classes de risque, sont bien du même ordre de grandeur.

professionnels, les outils habituellement utilisés pour noter les clients ou les produits (au moment de l'octroi dans ce cas) sont majoritairement basés sur des méthodes de scoring. Le *backtesting* d'un score peut s'articuler autour de trois phases essentielles.

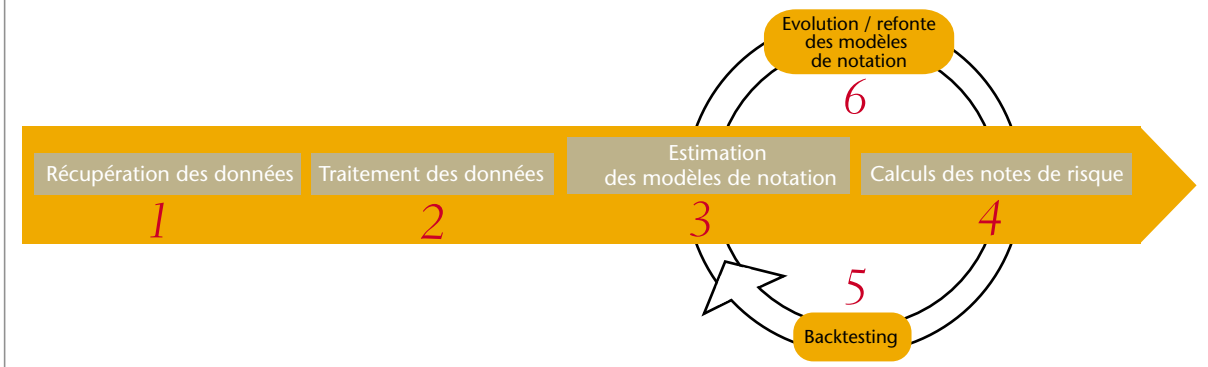
La première de ces phases consiste à s'assurer que le score en production est bien capable de prévoir le défaut. En d'autres termes, on s'assure par-là des capacités de discrimination du score ; la ROC

Curve (encadré ci-dessus) permet de réaliser ce test de façon graphique. En outre, l'indice de performance associé à cette courbe apporte un complément numérique à l'analyse graphique.

La seconde phase découle directement de la première. Elle a pour but de comparer les caractéristiques de la population actuellement scorée par rapport à celle qui existait au moment du développement du score. Cette analyse peut donc fournir des explications sur

une éventuelle dégradation des capacités de prévision du défaut. On s'assure par-là que les hypothèses en terme de structure de la population qui existaient au moment du développement du score sont toujours vérifiées : les performances du score, qui est un modèle statistique, reposent en effet sur ces hypothèses et il serait normal de constater une chute de ses performances si l'on observe un changement majeur de la structure de la population.

Schéma de construction et de maintenance d'un modèle de notation



La troisième phase réside dans la confrontation des taux de défaut prévus et constatés par classe de risque. Il s'agit de l'étape cruciale : c'est celle qui valide l'un des indicateurs majeurs du pilotage du risque, à savoir les probabilités de défaut (PD). En effet, pour décider de sa politique commerciale et du niveau de ses fonds propres, la banque utilise les PD sous-jacentes aux classes de risque (associées à un client ou un produit). Or, si ces classes de risque ne reflètent pas, a posteriori, les PD attendues alors cela signifie que le risque n'est plus sous contrôle.

- **Périodicité**

Le *backtesting* a entre autres pour objet de prévenir toute dégradation des performances des outils de notation, aussi, pour ce faire, il est nécessaire d'assurer un suivi quasi-permanent ou a minima un suivi récurrent de tous les outils de risque en production. Cette exigence permet d'anticiper très en amont toute dégradation des performances des outils et garantit ainsi un pilotage du risque rationnel et cohérent. Une revue trimestrielle détaillée paraît être une fréquence pertinente. Afin d'alléger la procédure, il serait également envisageable de produire un *backtesting* trimestriel synthétique (par exemple uniquement le suivi global des performances) enrichi d'une procédure annuelle plus complète.

Remarquons qu'un *backtesting* complet (performance de l'outil, stabilité de la population, comparaison des PD prévues et constatées) ne peut être mis en œuvre qu'après la survenance d'événements de défaut, ce qui n'est pas toujours adapté à un outil venant d'être mis en production. Néanmoins, il peut être intéressant de procéder à une vérification à chaud (quelques jours ou quelques semaines seulement après la mise en production), en s'assurant qu'il n'y a pas d'anomalie due au modèle de *score*, notamment en mettant en œuvre un test de stabilité. Ce test, dans le cas où il conclut à une stabilité de la population, n'implique pas obligatoirement de bonnes performances du *score*. En revanche, dans le cas contraire, il y a tout lieu de s'inquiéter quant aux capacités de discrimination de l'outil.

- **Compétences requises**

D'un point de vue pragmatique, le *backtesting* se compose de nombreux indicateurs déclinés en tableaux et autres graphiques. Son analyse requiert des compétences statistiques de bon niveau, afin d'exploiter au mieux toute la richesse de l'information et d'en comprendre les phénomènes sous-jacents, en particulier la dégradation de certains indicateurs, signe d'une diminution de la pertinence des modèles existants. Il ne s'agit pas seulement de relever les signaux d'alerte, encore faut-il en

identifier les causes et aussi les solutions à adopter. Par exemple, de nombreux effets saisonniers, complexes à appréhender, pourraient parfois laisser croire à une dégradation des performances des modèles, mais la pleine maîtrise de la méthodologie du *backtesting* et la capacité à distinguer la saisonnalité du comportement d'une population de celle du fonctionnement normal d'un outil sont le gage d'une appréciation objective de la situation. L'œil des experts-métiers est aussi un atout important pour expliquer ces phénomènes.

- **Coût**

Le coût d'un premier *backtesting* dépend à la fois du nombre et du type des outils de risque. Aussi, une solution destinée à amortir au plus vite les développements consiste à définir un protocole commun à tous les outils de risque (avec éventuellement des aménagements spécifiques selon les outils) : non seulement cette solution permet de gagner en temps de développement informatique mais aussi elle fournit plus de cohérence à l'ensemble du *backtesting*. Le coût de la méthodologie de *backtesting* et de sa mise en œuvre opérationnelle n'est pas négligeable, mais il est important d'investir suffisamment dans cette méthodologie censée rendre bâlo-compatible non seulement le fonctionnement des outils, mais aussi la communication concernant les outils de risque auprès de la Commission bancaire.

De façon à fiabiliser le *backtesting* et à minimiser son coût à la fois en termes de budget et de temps, il est indispensable de le réaliser via des procédures automatiques ou des progiciels dédiés. Ainsi chaque mise à jour ne nécessite qu'une intervention humaine mineure pour l'établissement des *reportings*. Le temps gagné sur la publication des *reportings* peut être consacré aux analyses, indispensables pour soulever des alertes et envisager des solutions ou au contraire pour signaler que les indicateurs de pertinence des modèles sont au vert.

Une fois cette étape d'industrialisation terminée, la banque se trouve alors dans un cercle vertueux : les modèles de notation, une fois construits et mis en production, sont périodiquement validés, ce qui conduit soit à les conserver en l'état, soit à les réestimer, voire à les refondre entièrement si la situation le nécessite (*graphique ci-contre*).

DU BACKTESTING AU STRESS-TESTING

Les analyses menées dans le cadre du *backtesting* constituent une bonne base pour prolonger vers d'autres études sollicitées par le Comité de Bâle : les scénarios de crise ou *stress-testing* (*encadré 2*).

Il s'agit ici de mesurer l'impact d'une variation prévisionnelle des paramètres PD/LGD/EAD sur le niveau de fonds propres. Cette mesure peut être considérée comme un indicateur de pilotage alliant à la fois des aspects risques et commerciaux : elle joue un rôle d'alerte dans le sens où elle indique le niveau des fonds propres en cas de dégradation du risque mais permet aussi d'informer les commerciaux de la consommation de fonds propres selon les stratégies qu'ils souhaitent développer, i.e. en fonction des EAD.

En outre, il est intéressant de projeter à un horizon de quelques mois à quelques années le porte-

2. Les scénarios de crise sollicités par le Comité de Bâle

« ...les banques doivent procéder à une simulation en regard du risque de crédit pour estimer l'effet de certaines conditions particulières sur leurs exigences de fonds propres réglementaires dans le cadre de l'approche NI. L'exercice est choisi par la banque et soumis à l'approbation de l'autorité de contrôle. Il doit être significatif et faire preuve d'une prudence raisonnable. Chaque établissement peut élaborer différentes méthodes applicables en fonction des circonstances. Il ne s'agit pas de demander aux banques de prévoir des situations catastrophiques, mais au moins d'envisager les effets de scénarios de légère récession, comportant par exemple deux trimestres consécutifs de croissance zéro, pour en déterminer l'incidence sur PD, PCD et ECD, etc. » (Extrait du texte réglementaire Bâle II d'avril 2003 -§ 397).

feuille de clients et d'encours actuel afin de saisir sa déformation en risque éventuelle, selon les hypothèses du moment. Par extension on peut alors analyser les impacts en termes de coût du risque et de fonds propres de nouvelles stratégies commerciales ou de nouvelles hypothèses conjoncturelles.

Exemples : Si l'on accroît de 5 % l'encours de crédit sur un certain segment de clientèle, quel est l'impact sur la consommation de fonds propres ou sur le coût du risque ? Si la conjoncture économique se dégrade l'année prochaine, voire si la croissance économique est nulle, quel est le risque supplémentaire induit ?

De tels scénarios, pour peu qu'ils soient réalistes, sont de nature à guider les décideurs ou à leur faire saisir la pleine dimension de leurs choix stratégiques et ainsi moduler ceux-ci, le cas échéant, en fonction des conséquences qu'ils peuvent avoir à court et moyen terme.

LES CLÉS DU SUCCÈS

Parmi les écueils à éviter, le plus important concerne la définition du périmètre des tests que l'on souhaite réaliser : quels sont les indicateurs à produire et à quelle fréquence ? En effet, les indicateurs qu'il est possible de calculer peuvent être très nombreux, aussi la

difficulté est de choisir ceux qui seront à la fois les plus utiles et les plus significatifs.

L'aspect système d'information revêt aussi une importance déterminante, pouvant à lui seul conditionner la faisabilité et la pertinence du *backtesting*, en fonction de la fiabilité et de la souplesse d'utilisation des informations qui le compose mais également de sa fréquence de mise à jour. Une information centralisée peut par exemple réduire considérablement les délais de production donc les coûts de développement des procédures de *backtesting*.

Il serait simpliste de considérer le *backtesting* comme un aspect technique particulier de la réforme de Bâle. Au contraire, celui-ci doit être vu comme un puissant liant entre les Systèmes d'Information, la méthodologie de calcul des PD et leur validation a posteriori, les acteurs risque et commerciaux, la communication auprès de la Commission bancaire. Envisager le *backtesting* de cette manière est la condition sine qua non pour le rendre utile, cohérent et pertinent et ainsi être en mesure de se justifier auprès des autorités compétentes.

En conclusion, la définition d'un protocole de *backtesting* et son modus operandi constituent un investissement non-négligeable néanmoins rentable au vu des enjeux qu'ils représentent. ■