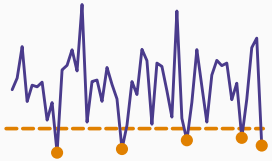
Mesure et gestion du risque de marché

Chapitre 6 : Les leçons de la littérature académique sur le trading book



Jaouad Madkour

Faculté d'Économie et de Gestion de Tanger · 30 août 2026

Plan du chapitre

Avant de commencer

Mettre en œuvre la VaR

Intégrer la liquidité

Choisir sa mesure de risque

Les tests de résistance

Unifiée ou compartimentée

La VaR, vue systémique

Synthèse

Avant de commencer

Pourquoi ce chapitre ?

Ce chapitre clôt le cours en prenant du recul. Il synthétise un rapport du comité de Bâle passant en revue la littérature académique sur la gestion du risque du trading book : l'horizon de VaR, la volatilité changeante, le backtesting, la liquidité, le choix de la mesure de risque, les tests de résistance, l'agrégation des risques, et — question devenue centrale depuis la crise de 2008 — le rôle procyclique de la VaR elle-même dans la stabilité du système financier.

Mettre en œuvre la VaR

Quel horizon retenir ?

Une question sans réponse évidente

Christoffersen et Diebold qualifient l'horizon optimal de VaR de « question évidente sans réponse évidente ». Le consensus académique est que l'horizon approprié dépend des caractéristiques de la position — sa nature, sa liquidité — et non d'une règle uniforme applicable à l'ensemble du portefeuille.

La critique de Danielsson

Si la VaR à dix jours vise à protéger contre une crise de liquidité, l'horizon réglementaire est mal calibré : au niveau de confiance de 99 %, un dépassement de la VaR à dix jours survient en moyenne 25 fois par décennie, alors qu'une véritable crise de liquidité systémique est un événement bien plus rare — probablement moins d'une fois par décennie. La probabilité implicite du modèle et la fréquence réelle de l'événement qu'il est censé capter ne coïncident pas.

Contourner le problème par le scaling

Plutôt que d'estimer directement la VaR sur un horizon long, la pratique courante calcule la VaR à un jour puis la met à l'échelle par la racine carrée du temps. Solution commode, mais qui ne résout pas la difficulté sous-jacente : elle repose sur des hypothèses restrictives, rarement vérifiées, sur le processus générateur des rendements.

Le scaling par la racine du temps est-il fiable ?

Des hypothèses fragiles

Le scaling par racine du temps n'est valide que si les facteurs de risque sont normaux, centrés, et indépendants et identiquement distribués. Ces conditions sont rarement vérifiées empiriquement, surtout à haute fréquence.

Deux biais opposés selon le processus

Si les facteurs de risque suivent un processus GARCH(1,1), le scaling par racine du temps **surestime** la volatilité et donc la VaR à long horizon. S'ils suivent un processus à sauts (jump-diffusion), il la **sous-estime** — le risque de queue est alors systématiquement biaisé à la baisse. Dans un cas comme dans l'autre, la simplicité du scaling se paie par une distorsion selon la vraie dynamique des rendements, et la littérature n'offre aucune alternative pleinement satisfaisante pour la VaR conditionnelle.

Faut-il intégrer une volatilité changeante ?

Trois méthodes industrielles

L'approche EWMA de RiskMetrics pondère les observations passées par un facteur de décroissance $\lambda = 0{,}97$ — un cas particulier contraint d'un modèle IGARCH(1,1). La pondération par l'âge de Boudoukh, Richardson et Whitelaw offre une alternative plus simple. La simulation historique filtrée (FHS), en filtrant d'abord les rendements par un modèle GARCH, s'est révélée plus sensible aux changements de volatilité — Pritsker a montré que la pondération BRW manque justement de cette sensibilité.

Le vrai obstacle : la dimension

Modéliser une volatilité et des corrélations changeantes devient rapidement impraticable à grande échelle : les modèles GARCH multivariés usuels (BEKK, DCC) sont difficiles à estimer dès que le nombre de facteurs de risque grandit. Des solutions récentes existent — moyenner les vraisemblances avant estimation, ou réduire la dimension par des modèles à facteurs dynamiques — mais aucune n'est pleinement satisfaisante pour de très grands portefeuilles.

Faut-il intégrer une volatilité changeante ? (suite)

Un arbitrage avec la procyclicité

Intégrer la volatilité changeante rend le modèle plus réaliste et réactif, mais introduit un risque de procyclicité : une VaR qui gonfle mécaniquement en période de stress peut amplifier, plutôt qu'amortir, les tensions de marché — thème qui reviendra en fin de chapitre.

Le backtesting, revisité

Les limites du simple comptage

Le test de Kupiec sur le nombre d'exceptions est simple mais souffre d'un manque de puissance chronique, en particulier avec peu d'observations (250 jours) ou à un niveau de confiance élevé (99 %) — configuration typique du contexte réglementaire.

Un problème d'horizon persistant

Les banques testent en pratique leur VaR à un jour, alors que la VaR réglementaire porte sur dix jours. Rien ne garantit qu'une VaR à un jour bien calibrée implique une VaR à dix jours également bien calibrée — et la littérature reste silencieuse sur cette question, pourtant cruciale.

Au-delà du comptage : la magnitude et la dynamique

Lopez propose de pondérer les tests par l'**ampleur** des dépassements, pas seulement leur nombre. Christoffersen, prolongé par Berkowitz, Christoffersen et Pelletier, teste en outre l'**indépendance temporelle** des exceptions — un modèle correctement calibré ne doit produire aucune autocorrélation dans ses dépassements.

Intégrer la liquidité

Deux formes de liquidité

Liquidité exogène et endogène

La liquidité **exogène** est le coût de transaction standard, pour une taille d'ordre usuelle, mesuré par l'écart bid/ask. La liquidité **endogène** est le coût additionnel provoqué par la taille même de la transaction, lorsqu'elle est assez importante pour déplacer le prix de marché — un coût que la transaction elle-même génère, plutôt que subit.

Intégrer la liquidité exogène

Bangia, Diebold, Schuermann et Stroughair proposent d'ajouter à la VaR un coût de liquidité fondé sur la variabilité de l'écart bid/ask pour des transactions de taille standard — un ajustement relativement simple, car il repose sur des données de marché directement observables.

Deux formes de liquidité (suite)

La liquidité endogène, un défi bien plus ardu

L'impact d'une liquidation de grande taille sur les prix n'est presque jamais intégré, ni dans la valorisation du portefeuille, ni dans la VaR — alors que son impact potentiel peut être considérable, en particulier pour les dérivés complexes des grandes institutions. Le couvrement dynamique d'options peut lui-même amplifier ce risque : un gérant qui achète quand le prix monte et vend quand il baisse, si sa position est significative par rapport au marché, aggrave le mouvement qu'il subit.

Liquidité et horizon de risque

La littérature suggère une interdépendance : lorsque la liquidité de marché se dégrade, l'horizon de risque approprié s'allonge — et l'exposition au risque de crédit tend elle-même à augmenter avec un horizon plus long. Liquidité et risque ne sont donc pas des dimensions indépendantes du problème.

Choisir sa mesure de risque

Ce que la VaR ignore

La VaR ne renseigne que sur le quantile — elle ne dit rien de l'ampleur des pertes au-delà. Un gestionnaire de risque strictement piloté par la VaR peut ainsi être incité à réduire les pertes juste en-deçà du seuil, tout en acceptant des pertes bien plus sévères au-delà — exactement l'inverse de l'intérêt des régulateurs, qui doivent absorber les pertes extrêmes.

La cohérence, rappel

Une mesure de risque est dite **cohérente** si elle vérifie la sous-additivité (la diversification ne peut qu'aider), l'homogénéité positive, et la monotonicité. La VaR peut violer la sous-additivité : deux positions dont chacune a une VaR nulle peuvent, combinées, afficher une VaR positive.

Les limites de la VaR (suite)

Un défaut surtout théorique ?

La VaR est sous-additive sous des hypothèses assez générales — notamment lorsque la distribution jointe des facteurs de risque est elliptique (la loi normale multivariée en est un cas). Les contre-exemples les plus frappants de non-sous-additivité restent artificiels ou reposent sur des distributions à moments infinis. Mais Degen, Embrechts et Lambrigger montrent que le problème redevient concret pour des indices de queue modérés, pertinents pour le trading book — la VaR à 99 % peut alors être **super-additive** même pour des distributions raisonnables.

Trois corrections apportées par l’ES

L’ES corrige trois défauts de la VaR : elle tient compte de la sévérité des pertes au-delà du seuil ; elle est toujours cohérente, donc toujours sous-additive ; et elle atténue l’arbitraire du choix d’un seuil de confiance unique, puisqu’elle intègre l’ensemble de la queue.

Le prix à payer

L’ES est plus coûteuse à calculer et moins robuste à backtester que la VaR : les tests usuels reposant sur le comptage des dépassements ne suffisent plus, il faut tenir compte de leur ampleur. Le débat empirique sur la précision comparée des deux mesures reste ouvert, mais les techniques de simulation avancées — notamment l’échantillonnage préférentiel — ont rendu le calcul de l’ES praticable en routine.

Les mesures spectrales

Généraliser l'ES par une pondération libre

Une mesure spectrale remplace la pondération uniforme de l'ES par une fonction de poids croissante sur la queue de la distribution — donnant plus de poids aux pertes les plus extrêmes qu'aux pertes juste au-delà du seuil. L'ES en est le cas particulier à poids constant ; la VaR en est, formellement, un cas limite dégénéré.

Peu utilisées en pratique, malgré leurs atouts

Les mesures spectrales restent rares dans l'industrie bancaire, malgré une meilleure régularité analytique et un lien plus intuitif avec l'aversion au risque. Le secteur de l'assurance utilise en revanche des concepts apparentés — les mesures de distorsion, dont la transformée de Wang est un exemple notable.

D'autres mesures, moins utilisées

La variance reste historiquement dominante mais exige un second moment fini et n'est adaptée qu'aux distributions approximativement symétriques — un défaut sérieux pour des pertes de trading book souvent asymétriques. La déviation moyenne, les moments partiels supérieurs et la mesure de queue gauche offrent des alternatives, chacune avec ses propres limites de cohérence ou de calcul.

Les tests de résistance

Trois familles de scénarios

Historiques, hypothétiques, ou mécaniques

Un test de résistance peut reposer sur un événement historique reproduit, sur un scénario hypothétique construit à dessein, ou sur une recherche mécanique automatisée du pire résultat sur un ensemble de facteurs de risque. Chaque approche explore la queue de la distribution des pertes au-delà du seuil habituel de la VaR — typiquement 99 %.

La faiblesse classique des tests de résistance

Sans modèle de risque sous-jacent, la probabilité de chaque scénario reste inconnue — impossible d'en évaluer l'importance relative, et rien ne garantit que les scénarios les plus plausibles et extrêmes aient été considérés. Intégrer les tests de résistance dans la modélisation probabiliste du risque, en assignant des probabilités aux scénarios, permet d'unifier ces deux exercices aujourd'hui souvent incompatibles.

La révision de juillet 2009

Depuis les révisions de juillet 2009 du cadre de Bâle II, les banques doivent calculer une VaR stressée : le même moteur de calcul habituel, mais calibré sur une période continue de douze mois de stress financier significatif pour le portefeuille de la banque.

Un remède partiel

La VaR stressée agit davantage « à travers le cycle » que la VaR ordinaire, atténuant partiellement la procyclicité — mais elle n'en constitue qu'une composante parmi d'autres du capital total requis, et sa mise en œuvre technique — corrélations stressées, distributions à queue épaisse — reste un champ encore peu exploré par la littérature académique au moment de ce rapport.

Unifiée ou compartimentée

L'approche par blocs

Le cadre bâlois actuel

Le capital réglementaire additionne, sans les faire interagir, les exigences calculées séparément pour le risque de marché, le risque de crédit et le risque opérationnel — une approche dite **compartimentée**, par opposition à une approche **unifiée**, qui modéliserait simultanément tous les risques en tenant compte de leurs interactions.

La tentation de la diversification

L'industrie estime les bénéfices de diversification ignorés par la simple addition entre 10 et 30 %. Mais cette intuition suppose implicitement qu'on puisse découper le portefeuille global de la banque en sous-portefeuilles purs, chacun exposé à un seul type de risque — une hypothèse rarement vérifiable en pratique.

Le risque de change croisé

Un prêt en devise étrangère illustre le problème : une dépréciation de la devise (risque de marché) augmente le fardeau de remboursement de l'emprunteur et donc sa probabilité de défaut (risque de crédit) — les deux risques sont ici mécaniquement liés, pas séparables. C'est le « wrong way risk », bien connu en risque de contrepartie, généralisé ici à l'interaction marché-crédit.

Compounding contre diversification : ce que montrent les études empiriques

Un indicateur simple pour comparer les études

La plupart des études rapportées définissent un « indice de diversification inter-risques » — le ratio entre la mesure de risque unifiée et la somme des mesures compartimentées. Un ratio inférieur à 1 signale une diversification réelle ; un ratio supérieur à 1 signale un **compounding**, une amplification du risque total par l'interaction entre catégories.

Des résultats contrastés selon l'approche

Les études « bottom-up », qui modélisent explicitement l'interaction entre facteurs, trouvent des ratios très variables — de 0,60 à près de 4,00 selon les portefeuilles et hypothèses, avec des cas avérés de compounding. Les études « top-down », qui supposent d'emblée la séparabilité des risques pour ensuite les recombiner par une matrice de corrélation, trouvent presque systématiquement de la diversification, avec des ratios de l'ordre de 0,42 à 0,90 — un résultat qui découle largement de l'hypothèse de départ, pas d'une preuve indépendante.

Compounding contre diversification : ce que montrent les études empiriques (suite)

La leçon centrale

La possibilité de former des sous-portefeuilles homogènes par type de risque est une condition **suffisante** pour que la diversification joue en faveur d'une approche compartimentée. Mais cette condition est rarement vérifiable dans un bilan bancaire réel, où marché et crédit s'entremêlent — d'où le risque, documenté empiriquement, que l'approche compartimentée sous-estime le capital réellement nécessaire.

La VaR, vue systémique

Le levier, miroir inversé de la valeur des actifs

Pour un bilan simplifié où la dette reste approximativement constante, le levier — le ratio actifs sur fonds propres — varie en sens inverse de la valeur de marché des actifs : quand les actifs prennent de la valeur, les fonds propres augmentent plus vite proportionnellement, et le levier baisse mécaniquement.

Rétablir un levier cible de 10

Un bilan de 100 en actifs, 90 en dette, 10 en fonds propres affiche un levier de 10. Si les actifs progressent de 1 (à dette inchangée), les fonds propres passent à 11 et le levier tombe à $101/11 \approx 9{,}18$. Pour ramener le levier à 10, la banque achète 9 d'actifs supplémentaires financés par 9 de dette nouvelle : le bilan atteint 110 d'actifs, 99 de dette, 11 de fonds propres, et $110/11 = 10$. La gestion active du bilan transforme ainsi une hausse de prix en achat supplémentaire — l'inverse du mécanisme stabilisateur habituel de l'offre et la demande.

Le mécanisme de la procyclicité (suite)

Un « accélérateur d'incendie » plutôt qu'un extincteur

Un système d'intermédiaires pilotés par une contrainte de VaR amplifie les mouvements de prix au lieu de les amortir : la hausse des prix relâche la contrainte de VaR et incite à acheter davantage, la baisse la resserre et force à vendre — précisément au moment où le marché est le moins liquide. Shin qualifie cette dynamique d'« accélérateur d'incendie », par contraste avec l'intention réglementaire d'« extincteur ».

La procyclicité du capital réglementaire

Si le capital exigé est proportionnel à la VaR par unité d'actif, et que cette VaR par unité d'actif est elle-même **contracyclique** — plus faible en période calme, plus élevée en période de stress —, alors le levier autorisé est mécaniquement **procyclique** : plus élevé quand les prix montent, plus faible quand ils chutent, exactement au mauvais moment.

Le mécanisme de la procyclicité (suite 2)

Une preuve empirique convergente

Adrian et Shin documentent une relation positive et robuste entre la taille du bilan et le levier des grandes banques d'investissement américaines avant la crise de 2008 — cohérente avec le mécanisme de gestion active du bilan décrit ci-dessus, et avec un effet mesurable sur les primes de risque et la volatilité agrégée des marchés financiers.

Que faire de ce constat ?

Un problème qui dépasse la sophistication du modèle

Le mécanisme amplificateur ne dépend pas de la sophistication du modèle de risque utilisé — il subsisterait à l'identique si la VaR était remplacée par l'expected shortfall, ou par une mesure de risque encore plus raffinée. C'est la mécanique du pilotage par une contrainte de risque calculée sur des données récentes, et non le choix de la mesure elle-même, qui produit la procyclicité.

Une piste, encore partielle

La VaR stressée de 2009 agit « à travers le cycle » et atténue partiellement le problème, mais elle ne représente qu'une fraction du capital total exigé. La littérature reste divisée sur la nécessité d'un abandon plus radical de l'approche fondée sur la VaR au profit d'une perspective véritablement systémique du risque endogène.

Synthèse

Ce qu'il faut retenir

Synthèse du chapitre

Sur l'**horizon** de VaR, la littérature ne trouve pas de règle universelle : il dépend de la position, et le scaling par racine du temps, bien que commode, repose sur des hypothèses rarement vérifiées. Sur la **volatilité changeante**, EWMA, pondération par âge et simulation historique filtrée offrent des solutions imparfaites, d'autant plus difficiles à grande échelle que le nombre de facteurs croît. Le **backtesting** manque de puissance au niveau de confiance et à la taille d'échantillon usuels en régulation, et doit intégrer non seulement le nombre mais aussi l'ampleur et la dynamique temporelle des dépassements. La **liquidité**, exogène ou endogène, reste largement absente des modèles de VaR malgré son importance démontrée durant les crises.

Ce qu'il faut retenir (suite)

Synthèse du chapitre (suite)

Sur le choix de la **mesure de risque**, l'expected shortfall corrige les défauts conceptuels de la VaR — non-coherence, insensibilité à la sévérité des pertes extrêmes — au prix d'un calcul et d'un backtesting plus exigeants ; les mesures spectrales généralisent encore cette logique, sans s'être imposées en pratique. L'agrégation **unifiée versus compartimentée** des risques dépend d'une hypothèse de séparabilité rarement vérifiée : la diversification n'est pas garantie, et des effets de compounding sont empiriquement documentés. Enfin, au niveau **systémique**, la gestion active du bilan sous contrainte de VaR transforme une technologie censée stabiliser le système en un mécanisme procyclique d'amplification des booms et des crises — un constat qui dépasse le choix de la mesure de risque et interroge le principe même du pilotage réglementaire par des modèles internes.