



HAL
open science

De l'usage des statistiques pour le pilotage et la régulation des risques en assurance : les limites de l'approche adoptée par Solvabilité 2

Sylvestre Frézal

► **To cite this version:**

Sylvestre Frézal. De l'usage des statistiques pour le pilotage et la régulation des risques en assurance : les limites de l'approche adoptée par Solvabilité 2. Gestion et management. Université de Strasbourg, 2020. Français. NNT : 2020STRAB009 . tel-03264302v2

HAL Id: tel-03264302

<https://theses.hal.science/tel-03264302v2>

Submitted on 18 Jun 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ÉCOLE DOCTORALE AUGUSTIN COURNOT
LaRGE

THÈSE présentée par :
Sylvestre FREZAL

soutenue le : 25 septembre 2020

pour obtenir le grade de : **Docteur de l'université de Strasbourg**

Discipline/ Spécialité : Sciences de gestion

**DE L'USAGE DES STATISTIQUES POUR LE PILOTAGE
ET LA REGULATION DES RISQUES EN ASSURANCE**
Les limites de l'approche adoptée par Solvabilité 2

THÈSE dirigée par :

Mme RAINELLI-WEISS Hélène

Mr ROGER Patrick

Professeur (Sc. de Gestion), Université de Rennes 1

Professeur (Finance), Université de Strasbourg

RAPPORTEURS :

Mme TADJEDDINE-FOURNEYRON Yamina Professeur (Sc. Economiques), Université de Lorraine

Mr ELIE Romuald

Professeur (Mathématiques appliquées), UMLV

AUTRES MEMBRES DU JURY :

Mr FRANCOIS Pierre

Mr WEILL Laurent

Directeur de recherche (Sociologie), CNRS - Sciences Po

Professeur (Sc Economiques), Université de Strasbourg

Sommaire synthétique

<i>Remerciements</i>	3
<i>Introduction Générale</i>	5
<i>Partie I : Les limites de l'approche statistique adoptée par Solvabilité 2</i>	32
<i>Chapitre 1 : Le problème de la fiabilité des estimateurs</i>	32
<i>Chapitre 2 : La question de la pertinence des indicateurs</i>	70
<i>Partie II : Dysfonctionnements des mesures de risque dans la régulation des assurances : proposition d'un cadre théorique et applications au coût du risque</i>	97
<i>Chapitre 3 : Les défauts de l'approche statistique des risques : proposition de théorisation et application à Solvabilité 2</i>	97
<i>Chapitre 4 : Coût et tarification des risques : de Solvabilité 2 à la question de l'équité des prix des assurances</i>	131
<i>Partie III : Retour sur les facteurs d'institutionnalisation de l'approche statistique promue par Solvabilité 2</i>	159
<i>Chapitre 5 : La réussite d'un amalgame problématique</i>	159
<i>Chapitre 6 : Le basculement vers une logique marchande</i>	202
<i>Conclusion Générale</i>	270
<i>Références Générales</i>	276
<i>Sommaire Détaillé</i>	284

Remerciements

Une thèse est un marqueur dans un parcours ; parcours fait de progressions et de bifurcations, de décalages et de retours... mouvements dont chacun est à la fois la résultante et la source de rencontres et d'échanges, d'approfondissements et d'ouvertures, d'opportunités et d'amitiés.

Sans les uns et les autres, cette thèse n'aurait ainsi pas existé ou, à tout le moins, aurait été différente dans son objet ou dans ses analyses. Nombre sont donc ceux qui auront eu une influence significative, structurante ou décisive sur cette thèse.

En premier lieu, je tiens à remercier Hélène Rainelli-Weiss, qui a accepté de la diriger et dont les conseils ont été particulièrement précieux, tant pour l'architecturer que pour parvenir à en mettre les différentes composantes en perspective ; qu'elle soit également remerciée pour son investissement tout aussi crucial dans le sprint final.

Je tiens également à remercier particulièrement Patrick Roger pour son soutien en co-direction, ainsi que Romuald Elie et Yamina Tadjeddine-Fourneyron qui ont accepté d'en être rapporteurs, qui plus est dans des circonstances géographiques et/ou temporelles exigeant un investissement tout particulier dans cette tâche. De chaleureux remerciements également à Laurent Weill, qui a bien voulu accepter de participer au jury.

Parmi les membres du jury, je tiens à rendre un hommage particulier à Pierre François. Pendant de longues années, il fut un interlocuteur, un ami, puis un compagnon de route. Je lui suis redevable non seulement de discussions passionnantes et de découvertes variées, mais aussi des travaux constituant cette thèse, certains directement et sans doute tous indirectement.

Au nombre de ceux sans qui ces travaux n'auraient certainement pas vu le jour, il faut citer, naturellement, Hélène N'diaye et Virak Nou pour leur soutien comme pour leurs idées séminales ou en rebond, ainsi que, bien sûr, Pierre Arnal, Pierre Le Moine et Emmanuel Sales ; Antoine Frachot et Benoît Ravel.

Que soient également remerciés tous ceux dont les discussions m'ont alimenté ou éclairé, pendant les travaux ou longtemps en amont : Nathanaël Abecera, Stéphane Dedeyan, Hélène Denis, Stéphan Fangue, Eléonore Haguët, Bertrand Labilloy, Tanguy Le Maire, Sandrine Lemery, Eric Lombard, Antoine Mantel, David Mariuzza, Pierre Michel, Carine Ollivier, Marc Porin, Arthur Silve et Fabrice Staad notamment.

Enfin, je suis également particulièrement reconnaissant à Laurence Barry, tant pour les échanges communs et travaux que nous avons menés ensemble que pour la continuité et la pérennité qu'ils lui doivent désormais.

Il va sans dire que cette liste ne saurait être exhaustive. Y manquent, parmi d'autres, toutes les personnes dont la présence a constitué un terreau propice à une telle aventure. Leur omniprésence rend vaine le fait de coucher leur nom sur le papier.

Introduction Générale

1. Positionnement du sujet

1.1 Contexte : l'assurance, un secteur crucial profondément impacté par la réforme prudentielle Solvabilité II

Le secteur de l'assurance est d'une importance économique et sociale considérable. Ainsi, les placements des assureurs français représentent 2 350 milliards d'euros (FFA, 2017), un montant supérieur au PIB : ils sont des investisseurs institutionnels dont l'impact sur le financement de l'Etat et des entreprises est décisif ; ce secteur représente, en France, environ 150 000 emplois directs (OEMA, 2017) ; enfin, et surtout, les services d'assurance, par les protections qu'ils fournissent, sont une source de sérénité, et partant de capacité de prise de risque des personnes tant physiques que morales, cruciale pour le dynamisme de nos sociétés. Malgré cette importance décisive, l'assurance est, tant dans l'absolu qu'au regard des marchés financiers et de la banque, extrêmement peu étudiée par les sciences sociales, qu'il s'agisse de l'économie, des sciences de gestion, ou de la sociologie (par exemple, le terme *insurance* n'apparaît que trois fois dans l'index du *Oxford handbook of the sociology of finance*).

Ce secteur est aujourd'hui percuté par de nombreuses ruptures (évolutions technologiques, contexte financier, etc.), à l'instar d'autres pans de notre économie. En Europe, il doit également faire face, et cela lui est davantage spécifique, à une profonde évolution de son cadre de régulation : une réforme du système prudentiel (Solvabilité 2, EC, 2009) a été conçue dans les années 2000 (François, 2015) et commence à être mise en œuvre. Indissociable d'un état d'esprit financier modelé par les mathématiques financières développées

dans la seconde moitié du vingtième siècle (cf. *infra*), cette révolution de la réglementation est susceptible d'engendrer des transformations très sensibles des équilibres du secteur.

De fait, les impacts de cette réforme sont considérables :

- directement, en termes financiers pour les compagnies d'assurance, à en croire la presse et les fédérations professionnelles : des coûts de mise en conformité initialement anticipés supérieurs au milliard d'euros (Thévenin & Vial, 2009), et plusieurs dizaines de milliards d'euros de capitaux propres potentiellement à lever, selon le contexte macro-financier ;
- indirectement, en termes de modification de l'industrie et du marché de l'assurance, et de répercussions économiques et sociales (voir par exemple Frunza, 2013). On peut ainsi tenir pour très probables les évolutions suivantes :
 - des fusions/euthanasies de petites structures dont la complexité permettait une gestion « artisanale » peu en adéquation avec les exigences de gouvernance structurée et d'industrialisation des métriques introduites par la directive, et la poursuite du regroupement de structures ayant intérêt à jouer sur les « bénéfices de diversification » pour optimiser leurs exigences de fonds propres globales ;
 - des changements d'allocation d'actifs des assureurs ayant des conséquences massives sur le financement de l'économie, avec par exemple une forte baisse des investissements en action (pénalisés par les calibrages spécifiques introduits par cette réforme) ;

- des changements dans l'offre de produits des assureurs : une épargne moins garantie, un arrêt de l'offre ou une hausse considérable des prix des "branches longues" (retraite, responsabilité civile, construction, etc.), du fait d'exigences de fonds propres plus fortes qui dégradent la rentabilité de ces produits.

La réforme dont il est question s'appuie sur la transformation très sensible de l'appréhension et la gestion des risques financiers, pour lesquelles les modèles et les mesures sont largement substitués aux heuristiques et intuitions qualitatives antérieures (Foulquier & Arias, 2012). Ainsi, *Solvabilité 2* instaure une nouvelle comptabilité fondée sur des modèles complexes issus de la finance de marché (Planchet & Thérond, 2007). En outre, elle modifie profondément les modalités de détermination du niveau minimum de fonds propres devant être détenus par les assureurs, incorporant ici encore des concepts issus des mathématiques financières. Alors qu'auparavant, le niveau minimum de fonds propres était relativement forfaitaire et fonction principalement de la taille de chaque compagnie (EC, 1992, et EPC, 2002), il devra désormais être déterminé en fonction des risques réputés pris par la compagnie : risque de baisse des marchés immobiliers si elle investit en immobilier, de faillite des émetteurs obligataires, de tempête, de pandémie, d'erreur dans les comptes, de mauvaise tarification, de rachat prématuré des assurés, de dérive des frais de gestion, etc. (Guibert et al., 2014). Pour rendre cette obligation opérationnelle, une mesure de risque est définie et des modèles extrêmement complexes sont développés, modèles ayant pour objectif de quantifier l'ensemble des risques susceptibles de peser sur une compagnie d'assurance et de les agréger en un seul chiffre. Cette refonte de la régulation, qui participe du même état d'esprit que les réformes comptables (normes IFRS) également menées dans l'industrie de l'assurance (Amenc et al, 2006), repose donc sur des outils dont les concepts épistémiques sous-jacents sont

précisément ceux introduits par l'industrie financière au cours des dernières décennies, et dont la crise financière de 2007 laisse penser qu'il serait utile de mieux en caractériser le périmètre de pertinence et les conséquences.

1.2 Etat de l'art : les statistiques, un outil de gestion des risques qui dysfonctionne

Les statistiques (espérance, moyenne, médiane, puis probabilité) ont émergé, jusqu'au 18^{ème} siècle, dans l'objectif d'éclairer la prise de décision en incertain. Le recours aux – et la construction des – statistiques, au service de cette ambition de rationalisation et d'optimisation des décisions, s'attachait alors à un objectif descriptif avant d'être normatif (décrire la façon dont les gens éclairés prennent leurs décisions, afin que ceux qui ne disposent pas du même intellect et de la même tempérance puissent les imiter). Toutefois, les désillusions (e.g. le paradoxe de Saint-Pétersbourg) et les échecs des tentatives d'amélioration successives (e.g. par Bernouilli) conduisirent, au 19^{ème} siècle, à récuser leur utilisation dans ce cadre, Mill et Bertrand la qualifiant de « scandale des mathématiques » (voir par exemple Porter, 1986, pour une histoire longue détaillée ou Daston, 1989, pour une présentation plus synthétique).

Au 20^{ème} siècle toutefois, l'usage des grandeurs statistiques pour éclairer la décision en situation de risque (Knight, 1921) fut réintroduit et s'est généralisé, dans la littérature économique, à partir de la formalisation de l'espérance d'utilité par Von Neumann & Morgenstern (1944) et son axiomatisation par Savage (1954). A l'orée de cette réintroduction, la pertinence d'une mobilisation des outils statistiques à cette fin fut questionnée, sous un angle épistémologique (von Mises, 1949) ou logique (Allais, 1953 ; Samuelson, 1963). Depuis lors, les limites d'une approche statistique pour caractériser la décision en incertain et optimiser la gestion des risques ont été abondamment documentées, tant sous un angle descriptif et empirique (e.g. Ellsberg, 1961 ;

Kahneman & Tversky, 1979), que sous un angle normatif (e.g. Gilboa et al., 2008 ; Gilboa, 2009), ainsi que par certains praticiens (Taleb, 2007, sur les allocations d'actif ; Hadlane & Madouros, 2012, sur la régulation ; Walter & Pracontal, 2009, ou Mandelbrot & Hudson, 2009, sur les caractéristiques des modèles sous-jacents).

Toutefois, dans la littérature académique économique, ces critiques transversales restent largement circonscrites à la sous-discipline des théoriciens de la décision, sans qu'elles aient conduit à une remise en cause de leur usage, ni même à un questionnement significatif, dans d'autres champs de l'économie (Camerer, 1989 ; Gollier et al., 2013). Et, dans le champ de la finance, les grandeurs statistiques se sont largement imposées comme outil privilégié de prise de décision en incertain depuis le développement de la théorie du portefeuille par Markowitz (1952) et Sharpe (1964) qui inaugurèrent des développements d'autant plus féconds qu'ils concordèrent avec l'essor et le perfectionnement des outils informatiques. Au-delà de ces méthodologies d'allocation d'actif, les grandeurs statistiques y sont largement utilisées dans un domaine connexe, celui du pricing d'option, depuis l'article séminal de Black & Scholes (1973). Dès lors, les outils financiers fondés sur les grandeurs statistiques ne cessèrent de se complexifier, de se multiplier, et leur usage de se répandre au sein de la sphère financière. Les métriques se diversifièrent, avec l'apparition de la VaR (voir par exemple Adamko et al., 2015, pour une brève histoire) puis de la T-VaR comme nouvelle mesure de risque. Au-delà des enjeux d'investissement (gestion de portefeuille et tarification de dérivés) et, formellement du moins, de pilotage des risques associés, les usages s'étendirent à la communication financière, tant en matière d'explicitation de la politique des risques (e.g. Axa, rapport annuel 1997, p99) que de valorisation et de régulation financière, tout d'abord en banque (en matière de normes comptables et de régulation bancaires, voir notamment Chiapello, 2015 et

Baud & Chiapello, 2017), et plus récemment en assurance *via* la directive Solvabilité 2 que nous avons évoquée plus haut.

Pourtant, en dépit de la dissémination très importante de ces méthodes, de nombreux dysfonctionnements ponctuels liés à la mobilisation de ces outils pour la gestion et la régulation des risques ont été mis en lumière :

- d'une part, les difficultés de calibrage soulignées par la théorie (voir, pour une vision analytique, Eagle, 2004, ou, pour une vision historique consolidée, McGoun, 1995) se traduisent par des mesures de risque entachées d'une marge d'erreur technique très élevée,
 - qu'il s'agisse des mesures de risque, dans la continuité de celles mobilisées par la théorie du portefeuille et ses avatars (e.g. en banque, Danielsson, 2002 et 2008 ; en assurance, Kamega & Planchet, 2013) ;
 - ou des mesures bilancielle de valorisation, dans le prolongement de celles mobilisées par les théories de pricing d'option (El Karoui et al, 2015).

L'amplitude de ces marges d'erreur est telle qu'elle est susceptible de ruiner la pertinence de ces outils.

- D'autre part, l'usage de grandeurs statistiques dans ce cadre est porteur d'effets pervers :
 - d'un point de vue économique, car placer des mesures de risque quantitatives au cœur de la régulation est source de procyclicité (e.g. Kashyap & Stein, 2004 ; Taylor & Goodhart, 2004),
 - et, d'un point de vue organisationnel, car faire reposer la gestion des risques sur des grandeurs quantitatives conduit à une déresponsabilisation des acteurs (Mikes, 2009 et 2011 ; Power, 2009).

Comment expliquer que, malgré ces dysfonctionnements largement documentés, les statistiques demeurent aussi largement utilisées comme outils de décision en incertain, et notamment de gestion et de régulation des risques dans le secteur financier ?

Une histoire longue des statistiques (Daston, 1988 ; Porter, op.cit. ; Armatte, 2010), permet d'éclairer le contexte de leur émergence et les débats qui émaillèrent leur histoire. Elle permet de mettre en évidence l'importance des vertus pratiques (e.g. convaincre, et au-delà, trancher un désaccord par la comparaison de deux chiffres et partant mettre fin à une discussion) dans le recours et le choix des indicateurs statistiques.

Walter (2013) identifie trois natures de préoccupation ayant présidé à l'élaboration lente, à partir de la fin du 19^{ème} siècle, des outils stochastiques mobilisés en finance : morale, scientifique et d'efficacité opérationnelle. L'incorporation des outils mathématiques au sein du secteur financier, à compter de la seconde moitié du 20^{ème} siècle, a été documenté dans deux domaines : celui des marchés financiers (tarification d'options ; e.g. MacKenzie & Millo, 2003), et celui de la banque (comptabilité prudentielle ; e.g. Chiapello, op.cit.). On y retrouve deux des trois dimensions mentionnées *supra* se combinant pour expliquer la dissémination progressive de ces outils : un cadrage cognitif, découlant d'une période d'appropriation des concepts et des méthodes par les acteurs ; et un ensemble de vertus pratiques – de communication partagée, de justification, etc., vertus qui gravitent donc toujours, comme deux siècles auparavant, autour de cette capacité des chiffres à clore formellement un débat et permettre ainsi aux acteurs d'aller de l'avant. Si, malgré le cadrage cognitif, le caractère analytiquement non pertinent de ces outils est patent, les vertus pratiques peuvent être suffisamment décisives pour

emporter l'adhésion des acteurs (voir par exemple Mac Kenzie & Millo, 2009, sur la valorisation des appels de marge au sein des marchés d'option).

Ainsi, les outils des mathématiques financières se révèlent donc être, parfois, des outils de coordination plus que des outils d'analyse. Dès lors, si le développement de l'usage des outils statistiques associés aux mathématiques financières est indissociable de celui des marchés financiers, ceci peut ne pas être lu uniquement comme le fait que l'essor des marchés ait permis de mobiliser davantage de ressources concourant au progrès des outils statistiques mais, au contraire, comme le fait que le progrès des outils statistiques était l'une des conditions nécessaires à l'essor des marchés. Notamment, une des conditions nécessaires de développement d'un marché est la capacité de commensuration des titres (voir Espeland & Stevens, 1998, et Callon & Muniesa, 2005, pour un cadre général de la commensuration et de sa mise en œuvre par les marchés financiers ; Huault & Rainelli, 2011, sur le cas des dérivés climatiques ; et Huault & Rainelli, 2016, sur l'articulation entre cadres institutionnels et systèmes de commensuration, *via* l'exemple des marchés de gré à gré). Ainsi, une des sources de développement de l'usage des grandeurs statistiques peut être l'ambition de créer un marché.

Dans un tel contexte, la réglementation peut être un moteur de l'extension de l'usage de ces outils (Roe, 1994, Huault & Rainelli, 2007 et 2009, concernant les dérivés de crédit). Moteur parfois volontaire, lorsque, suivant la classification proposée par Boltanski & Thévenot (2006), la logique marchande parvient à fusionner avec la logique civique et à l'englober (Baud & Chiapello, *op.cit.*, ou Münzer, 2016, concernant la régulation prudentielle bancaire).

1.3 Question de recherche générale

Comme nous l'évoquions, la directive Solvabilité 2, après une quinzaine d'années de gestation, est entrée en vigueur début 2016, refondant la régulation prudentielle de ce secteur en Europe. Dans un souci de moderniser et d'améliorer la gestion et la régulation des risques des acteurs du secteur, elle s'appuie massivement sur les outils statistiques, les plaçant au cœur de son ambition d'être « *risk based* ». Ces grandeurs se retrouvent, de façon particulièrement explicite, dans les exigences de capital, où des paramètres jusqu'alors forfaitaires sont désormais réputés refléter une mesure de risque fine – une VaR, à l'instar de ce qui avait été introduit auparavant en banque, quoique suivant des paramètres et des modalités drastiquement différents. Elles se retrouvent également dans le bilan prudentiel, pour la détermination des passifs et, partant, des fonds propres disponibles figurant en regard de l'exigence de capital – via la notion de *best estimate*, en pratique une espérance.

Au vu des dysfonctionnements mis en lumière par la littérature économique, actuarielle, et en sciences de gestion des risques (cf. supra), cette extension du champ d'application de tels outils à la régulation des assurances interpelle. Ce domaine serait-il épargné par ces dysfonctionnements ? Qui plus est, la conception de Solvabilité 2 s'est pour partie déroulée, et a été finalisée, après la crise financière de 2008, qui aurait pu légitimement conduire à interrompre la généralisation de tels outils de pilotage des risques. Comment expliquer qu'ils aient continué de percoler vers ce secteur ?

Cette thèse (i) étend empiriquement au secteur de l'assurance la littérature produite dans d'autres secteurs afin d'y identifier les limites de l'utilisation des statistiques pour la gestion des risques et la régulation prudentielle, (ii) propose un cadre théorique unifié pour les dysfonctionnements liés à ces usages, et (iii) éclaire les ressorts de l'adoption de ces outils au sein du secteur de l'assurance.

Elle est donc composée de trois parties, dont les questions de recherche sont précisées *infra*, avant une présentation synthétique des résultats obtenus.

2. Questions de recherche

2.1 Les limites de l'approche statistique adoptée par Solvabilité 2

La première partie, composée de deux essais, montre que les dysfonctionnements associés aux mesures de risques développées aux fins de régulation prudentielle et de gestion des risques, bien documentés dans le secteur bancaire, se retrouvent également dans le secteur de l'assurance *via* la Directive Solvabilité 2. Le premier essai porte sur la pertinence des estimateurs, le second sur celle des indicateurs.

2.1.1 Le problème de la fiabilité des estimateurs

Ce premier essai¹ s'attache à caractériser l'amplitude des marges d'erreur associées aux mesures des risques de souscription non vie proposées par le régulateur européen à l'occasion des études d'impact successives ayant précédé la mise en œuvre de Solvabilité 2. Jusqu'à présent, les marges d'erreur des grandeurs prudentielles avaient été étudiées en banque d'un point de vue empirique (Harrison & Samuel, 2011, à partir des données du régulateur britannique), ou d'un point de vue théorique en jouant sur quelques paramètres (Danielsson, *op.cit.*) ; et en assurance de façon théorique, en jouant sur un unique type de paramètre : la volatilité alimentant la mesure des fonds propres chez El Karoui et al (*op.cit.*), la table de mortalité chez Kamega & Planchet (*op.cit.*), ou encore la méthode d'agrégation (e.g. Pfeifer & Strassburger, 2008).

¹ Frezal, S., 2018, Solvency II Is Not Risk-Based—Could It be? Evidence from Non-Life Calibrations », *North American Actuarial Journal*, 22(1), 15p.

Toujours², il s'agissait d'appréciations du risque exclusivement *techniques* (erreur dans le système de mesure scientifique), *externes* (indépendante des organisations qui utilisent ces grandeurs), *ponctuelles* (un unique ou un nombre très limité de paramètres d'entrée variant) et *absolues* (indépendante de l'usage opérationnel auquel les mesures étaient destinées). Mon analyse, plutôt que de mesurer la marge d'erreur absolue liée à un paramètre technique en faisant varier des paramètres d'entrée pour mesurer l'écart qui en résulte en sortie, étudie la dispersion de différents résultats de sortie fournis par le régulateur (EIOPA). En s'appuyant sur des données de panel, elle développe une méthodologie de ratio signal sur bruit (Sims, 2003) qui permet d'apprécier, plutôt que la fiabilité intrinsèque, la cohérence des différentes mesures entre elles. Contrairement aux travaux déjà existant, cette méthodologie ne permettrait pas de confirmer, le cas échéant, la fiabilité des mesures ; en revanche elle permet de l'infirmier en incorporant à la mesure de l'incertitude des sources de bruit autres que strictement techniques (e.g. politiques). Au-delà, j'exploite dans ce chapitre les données fournies par l'EIOPA pour justement caractériser les amplitudes relatives des différentes sources de bruit polluant les mesures de risque.

2.1.2 La question de la pertinence des indicateurs

Le deuxième essai³ porte sur la procyclicité de Solvabilité 2. Jusqu'à présent, la procyclicité des régulations prudentielles avait été étudiée dans le domaine bancaire, d'une part en ce qui concerne les fonds propres et l'impact de leur valorisation en valeur de marché (Plantin et al., 2008) et, d'autre part, en ce qui concerne les exigences de capital et l'impact de les déterminer de façon dite

² A l'exception, dans une certaine mesure, de l'étude menée à partir des données recueillies par le régulateur britannique (FSA), qui par exemple embarquait potentiellement les sources de variation non techniques, sans toutefois les distinguer les unes des autres.

³ Frezal, S., Haguet, E., and Nou, V., 2016, With-Profit Contracts Regulation and Procyclicality: A Comparative Analysis of Alternative Approaches, *Working paper PARI*, 31p., actuellement soumis au *Bulletin Français d'Actuariat* (demande de révisions)

risk-based (Kashyap & Stein, op.cit. ; Taylor & Goodheart, op.cit.). Notre étude développe un modèle qui permet d'appréhender, de façon globale, la procyclicité de la régulation sur un indicateur de pilotage agrégé. Surtout, elle porte sur l'assurance vie, secteur qui comporte un mécanisme contractuel susceptible d'accroître la procyclicité, au-delà des phénomènes déjà mis en lumière : nous modélisons ce mécanisme, en tirons des faits saillants théoriques, et confirmons par une simulation empirique le caractère significatif du mécanisme mis en lumière. Ceci permet enfin de caractériser une limite d'un indicateur de risque de type « VaR annuelle », en interrogeant la stabilité des incitations fournies au regard de la durée des risques portés.

2.2 Dysfonctionnements des mesures de risque dans la régulation des assurances : proposition d'un cadre théorique et applications au coût du risque
La deuxième partie, composée de deux essais, s'efforce d'apporter un fondement théorique transversal aux dysfonctionnements observés dans différents domaines et de déployer ce cadre à divers champs de la régulation des assurances.

2.2.1 Les défauts de l'approche statistique des risques : proposition de théorisation et application à Solvabilité 2

Pour ce faire, le premier essai⁴ propose une prise de recul en construisant un cadre conceptuel qui permet de distinguer les domaines où les outils statistiques améliorent la prise de décision de ceux où tel n'est pas le cas, avant d'appliquer ce cadre aux métriques sur lesquelles repose Solvabilité 2.

Cet essai s'inspire de la proposition de Doherty (2013) d'endogénéiser la description du risque : je n'y considère pas les mécanismes aléatoires en tant que tels, mais les rend indissociables de la position de l'observateur qui cherche

⁴ Frezal, S., 2016, Risk and statistics: a regulation fallacy, *Working paper PARI*, 35p.

à les décrire (et de l'utilisateur qui souhaite être éclairé par cette description). Ceci permet d'identifier deux situations polaires, et de lister les conséquences associées à l'usage de ces outils dans le type de situation où cet usage n'est pas pertinent. Ainsi, ce cadrage théorique permet d'unifier les différentes observations critiques associées à l'usage de grandeurs statistiques en matière de gestion des risques et de prise de décision dans l'incertain que nous avons rapidement synthétisées au début de cette introduction. Enfin, ce cadre théorique est appliqué à la régulation prudentielle Solvabilité 2, pour en décliner les conséquences opérationnelles dans ce secteur de l'assurance.

2.2.2 Coût et tarification des risques : de Solvabilité 2 à la question de l'équité des prix des assurances

Le second essai⁵ capitalise sur ce cadrage théorique pour explorer un champ connexe : celui de la régulation tarifaire. De fait, les régulations spécifiques à l'assurance ne se limitent pas à la question prudentielle de savoir quelles contraintes imposer aux assureurs pour limiter le risque, ou les impacts, d'une défaillance de leur part, mais portent également sur la légitimité de leurs tarifs et, en particulier, sur la légitimité de tarifs différenciés ou non selon les assurés (EC, 2004 & 2011, CJEU, 2011, CFA, 2013, Jones, 2013). Dans ce cadre, un point focal du débat est le positionnement du tarif individuel relativement à la prime pure (Daston, op. cit., Feller, 1968, Miller, 2009), c'est-à-dire le choix de considérer une grandeur statistique (l'espérance) comme point de référence approprié pour apprécier le service rendu à un acteur (l'assuré) confronté à une situation aléatoire.

L'apport de l'endogénéisation de la description du risque développée dans le premier essai trouve donc ici également à s'appliquer : par la distinction entre la position de l'assureur de celle de l'assuré lors de la souscription d'un contrat, il

⁵ Frezal, S. et Barry, L., 2019, Fairness in Uncertainty: some Limits and Misinterpretations of Actuarial Fairness, *Journal of Business Ethics* (2019) <https://doi.org/10.1007/s10551-019-04171-2>.

permet de questionner la pertinence de la transposition d'un outil statistique ayant fait ses preuves dans un champ – la moyenne comme moyen de tarifier un contrat d'assurance pour l'assureur, à un autre champ : la moyenne comme moyen de fixer un prix légitime pour l'assuré.

2.3 Retour sur les facteurs d'institutionnalisation de l'approche statistique promue par Solvabilité 2

La troisième partie, composée de deux essais, s'interroge sur les déterminants de la colonisation du secteur de la régulation prudentielle des assurances par ces outils statistiques.

2.3.1 La réussite d'un amalgame problématique

Le premier essai⁶ s'attache à déterminer si les deux principales sources de propagation des mathématiques financières identifiées au sein d'autres pans du secteur financier – un cadrage cognitif et des vertus pratiques – se retrouvent dans le secteur de l'assurance, jusqu'à présent largement délaissé par les *social studies of finance*.

Pour ce faire, nous nous appuyons sur deux cadres conceptuels. D'un point de vue méthodologique, nous menons notre analyse *via* le prisme du concept d'institution (François, 2011) comme cadre *a priori* (non susceptible d'être remis en cause, à l'instar de la danse de la pluie chez Hubert et Mauss, 1985), normatif (il oriente les décisions), et partagé (par l'ensemble des acteurs). C'est ainsi les modalités de l'institutionnalisation, dans cette acception, de ces outils – et de l'amalgame qui est en leur cœur – au sein du secteur de l'assurance que nous nous efforçons de caractériser. D'un point de vue épistémologique, nous nous appuyons sur le cadre développé dans la partie précédente où j'avais mis en évidence le fait que les statistiques sont mobilisées dans deux situations polaires, une où cela est pertinent et une autre où cela ne l'est pas. Nous

⁶ François, P., et Frezal, S., 2018, Instituer l'incohérence : aléa et hétérogénéité au sein du secteur assurantiel, *Sociologie du travail*, Vol. 60 – n° 1.

émettons l'hypothèse que le cadrage cognitif ayant permis à ces outils de diffuser malgré leurs dysfonctionnements correspond à un amalgame entre ces deux situations. Retracer la façon dont cet amalgame s'est progressivement imposé dans le secteur de l'assurance permet donc de documenter le cadrage cognitif comme l'une des sources de légitimité de ces outils dans ce secteur.

Au-delà, nous nous efforçons de documenter empiriquement les vertus pratiques dont cet amalgame peut être porteur et qui ont pu également contribuer à asseoir son usage. Ceci nous conduit à décrire la diversité du recul réflexif des différents utilisateurs vis-à-vis de ces outils.

2.3.2 Le basculement vers une logique marchande

Le second essai⁷ s'efforce de déterminer si, dans l'élaboration de la réglementation, la logique civique (protection des assurés) a pu servir de véhicule à la logique marchande (création d'un marché des passifs d'assurance), faisant de la poursuite de cette dernière le mobile essentiel de diffusion de ces outils. De fait, la volonté de commensuration des passifs d'assurance et des rentabilités des différentes *lines of business* pourrait être un motif clé de la généralisation de ces outils *via* la réglementation.

Une telle hypothèse peut être suggérée par des indices de plusieurs natures :

- en premier lieu, organisationnelle : la cellule assurance en charge de la directive Solvabilité 2 au sein de la Commission Européenne est rattachée à la DG Marché Intérieur ;
- en second lieu, juridique : le développement d'un marché – sans qu'il fût précisé lequel, potentiellement celui des passifs d'assurance – a été explicité au sein de la directive comme l'un de ses objectifs – quoique secondaire –. Or, il a été montré dans le domaine de la régulation

⁷ Frezal, S., 2015, De quoi Solvabilité 2 est-il le nom ?, *Working paper PARI*, 56p., devrait paraître comme chapitre au sein d'un ouvrage collectif aux Presses de Sciences Po.

bancaire que la priorité de construction d'un marché avait pu prendre le pas sur les objectifs prudentiels (e.g. Münzer, op.cit.) ;

- en troisième lieu, culturelle (dont il a été montré qu'il pouvait s'agir d'une barrière importante à l'émergence d'un marché, e.g. Zelizer, 1979) : les outils statistiques portés par Solvabilité 2 placent en leur cœur un changement de paradigme – la bascule d'une logique prudentielle où les engagements doivent pouvoir être *tenu*s à une logique prudentielle où les engagements doivent pouvoir être *repris*.

En ce sens, la mise en œuvre de ces outils participerait d'une moralisation de la logique marchande *via* la logique prudentielle. Afin de confirmer ou d'infirmier l'hypothèse reposant sur ce faisceau d'indices, je pars du principe que les choix techniques ne sont pas dénués de perspective politique et qu'il est donc nécessaire d'« *ouvrir la boîte noire [que constituent ces outils afin de] comprendre et décrire leur structure pour questionner l'évidence, [...] examiner et discuter les savoirs incorporés dans ces outils* » (Chiapello & Gilbert, 2009). Ce chapitre analyse donc les différentes métriques quantitatives de Solvabilité 2 incorporant ces outils statistiques (fonds propres, provisions et leurs différentes composantes, exigences de capital), afin de mieux comprendre si les solutions techniques retenues apportent des réponses à des questionnements prudentiels, ou bien marchands.

3 Résumé

3.1 Partie I : Les limites de l'approche statistique adoptée par Solvabilité 2

3.1.1 Chapitre 1 : Le problème de la fiabilité des estimateurs

“Solvency II Is Not Risk-Based. Could It Be? Evidence From Non-Life Calibrations.”⁸

Dans cet essai, je m'intéresse aux différents calibrages proposés par le régulateur européen (l'EIOPA) des risques de primes et de provisions (qui représentent conjointement environ la moitié des exigences de capital des organismes d'assurance non vie). Ces données sont disponibles pour différentes études d'impact successives ayant progressivement amené à déterminer le calibrage réglementaire et, pour chacune d'entre elles, pour différentes lignes d'activité. Intuitivement, la dispersion, au sein d'une étude donnée, des différentes mesures de risque associées à chacune des lignes d'activité correspond à l'amplitude du signal que la réglementation cherche à capter (telle activité est-elle plus risquée que telle autre, et devrait-elle donc mobiliser plus de capital ?). Réciproquement, la dispersion, pour une ligne d'activité donnée, des mesures de risque estimées au cours des différentes études correspond à la dispersion de plusieurs estimations d'une même grandeur – grandeur réputée stable dans le temps, puisque cristallisée dans la réglementation : cette dispersion correspond donc à l'amplitude du bruit qui vient polluer cette mesure.

Dans un premier temps, je m'assure que cette interprétation est correcte ; puis, je formalise plus finement la caractérisation relative de ces deux amplitudes ; enfin, ceci me permet d'apprécier, sur les données recueillies, la part relative de l'information fournie (quelle branche est effectivement la plus risquée, et dans quelle proportion ?) et celle du bruit qui la floute. Le principal résultat est que le signal est totalement noyé dans le bruit : ces mesures ne peuvent pas fonder, de façon fiable, une décision opérationnelle de pilotage des risques. Par ailleurs,

⁸ Frezal, S., 2018, Solvency II Is Not Risk-Based—Could It be? Evidence from Non-Life Calibrations », *North American Actuarial Journal*, 22(1), 15p.

une analyse comparable menée sur les calibrages du modèle interne d'une compagnie conduit à une conclusion analogue.

Dans la seconde partie de cet essai, je m'interroge sur les causes de ce bruit. Trois hypothèses sont successivement analysées, à la lumière des éléments fournis par l'EIOPA pour étayer ses calibrages. La première est celle d'interférences politiques ayant conduit à distordre le calibrage technique idéal, potentiellement de façon instable dans le temps. Si les données fournies par l'EIOPA ne permettent pas de conclure quant à la stabilité temporelle d'une telle distorsion, elles montrent néanmoins, et sans le moindre doute, que cette dimension politique est venue biaiser, de façon significative et systématique, les appréciations techniques de l'EIOPA. La seconde est celle d'une idiosyncrasie : les différentes compagnies et/ou les différents marchés d'assurance européens seraient trop disparates pour qu'un calibrage fondés sur certains d'entre eux (ou un calibrage moyen) apporte une information utile pour piloter les risques de telle ou telle autre compagnie (ou de tel ou tel marché). Ici encore, cette hypothèse est validée par les données : la grille d'analyse des risques « par ligne d'activité » retenue dans la directive ne semble pas permettre de construire des classes suffisamment homogènes pour que ces mesures permettent de caractériser de façon opérationnellement fiable les risques pesant sur une compagnie. La troisième hypothèse est celle d'une impasse technique : la diversité des choix de calibrage légitimes serait telle que les marges d'erreurs de telles mesures seraient tout à la fois considérables et incompressibles. Ici encore, les éléments fournis par l'EIOPA confirment cette hypothèse. Ceci, en un sens, confirme et étend les travaux menés sur les mesures de risques bancaires par exemple mais, par ailleurs, ne peut que questionner la pertinence de la philosophie sur laquelle repose Solvabilité 2.

3.1.2 Chapitre 2 : La question de la pertinence des estimateurs

“*With-Profit Contracts Regulation and Procyclicality: A Comparative Analysis of Alternative Approaches*”⁹

Fin 2011, dans un contexte de baisse des taux particulièrement douloureux pour les assureurs, les exercices d’impact menés dans la perspective de l’entrée en vigueur de Solvabilité 2 ont matérialisé une procyclicité des exigences quantitatives qui a surpris l’ensemble des acteurs par son ampleur. La procyclicité d’une réglementation dont les exigences de capital sont *risk based* (et donc s’accroissent en période de crise, lorsque les volatilités augmentent) et le bilan en valeur de marché (et donc dont les fonds propres diminuent immédiatement en période de crise, lorsque la valeur des actifs baisse) avait déjà été mise en évidence dans le domaine bancaire et ne constitue donc pas une surprise. Nous montrons que si son ampleur a surpris, c’est en raison de la combinaison de l’extension de la valeur de marché aux passifs (une nouveauté de ce cadre prudentiel) et d’un mécanisme contractuel propre à l’assurance-vie, celui de la « participation aux bénéfices ». Intuitivement, l’articulation est la suivante :

- avec des fonds propres en valeur de marché,
 - les perspectives de profits futurs des compagnies d’assurance sont valorisées dans le capital disponible ;
 - et les exigences de capital, correspondant à un risque de perte en capital, s’appuient donc sur une mesure des variations de profits futurs ;
- *via* le mécanisme de la « participation aux bénéfices »,
 - l’assureur ne conserve, en période de croisière, qu’une portion congrue des bénéfices (entre 0 et 15%), cédant le reste aux

⁹ Frezal, S., Haguët, E., and Nou, V., 2016, *With-Profit Contracts Regulation and Procyclicality: A Comparative Analysis of Alternative Approaches*, *Working paper PARI*, 31p.

assurés : si la situation se dégrade, ses résultats futurs – et partant ses fonds propres – varient peu puisque ce sont les assurés qui absorbent la majeure partie de la variation. Donc, l'exigence de capital est relativement faible ;

- mais en période de crise, l'assureur devra quoi qu'il en soit servir les taux garantis : si la situation se dégrade, ses résultats futurs en subissent les conséquences de plein fouet. Donc l'exigence de capital est relativement élevée.

Ce mécanisme, spécifique à l'assurance vie et à la façon dont les principes de Solvabilité 2 ont été mis en œuvre, s'ajoute aux autres sources de procyclicité déjà identifiées. Nous proposons un modèle théorique permettant de rendre compte de ce phénomène, ainsi qu'une simulation empirique qui confirme le caractère significatif des mécanismes ainsi mis en lumière.

Au-delà, nous mettons ainsi en évidence le fait que les exigences de capital ainsi construites n'envoient pas nécessairement les bons signaux. Ainsi, notre simulation empirique montre que si Solvabilité 2 avait été en vigueur au début des années 2000 (lorsque les taux d'intérêt étaient relativement élevés), son système de mesures agrégées des risques (bilan et exigences de capital) aurait conduit à inciter fortement les assureurs à augmenter leur activité retraite. De fait, les fonds propres créés par l'activité étant alors supérieurs aux exigences de capital engendrées par cette activité, le message envoyé par ces indicateurs aurait été : « plus on développe cette activité, moins on est risqué ». Au début des années 2010, le message aurait été exactement l'inverse : les fonds propres créés (perspectives de profits futurs) auraient été négatifs et, qui plus est, les exigences de capital auraient été considérablement accrues. Ainsi, le message envoyé aurait été : « cette activité est la plus risquée qui soit », et ce alors que seulement 10 années se sont écoulées entre temps, en regard d'une durée de ces produits largement supérieure à 20 ans.

3.2 Partie II : Dysfonctionnements des mesures de risque dans la régulation des assurances : proposition d'un cadre théorique et applications au coût du risque

3.2.1 Chapitre 3 : Les défauts de l'approche statistique des risques : proposition de théorisation et application à Solvabilité 2

*"Risk and statistics : the prudential regulation fallacy"*¹⁰

La première partie de cet essai formalise le champ de pertinence des statistiques dans la prise de décision. J'identifie ainsi deux cas polaires : les situations aléa, et les situations plurielles :

- si un décideur fait face à un événement aléatoire qui ne va se produire, dans le cadre de sa prise de décision, qu'une [ou peu de] fois, alors il est en situation aléa [l'aléa est effectivement significatif];
- si, dans le cadre de sa prise de décision, cet événement va se produire une infinité [ou un grand nombre] de fois, alors le décideur est en situation plurielle [l'aléa résiduel est négligeable en raison de la loi des grands nombres].

En situation d'aléa, le recours à des grandeurs statistiques repose sur un amalgame entre aléa et pluralité. Les principales conséquences en sont les suivantes :

- (1) En situation d'aléa, il n'est pas pertinent de fonder une décision sur la base de grandeurs statistiques associées à la fonction de distribution anticipée *ex ante*.
- (2) En outre, une analyse reposant sur de tels outils modifierait alors la perception du phénomène par le décideur, lui donnant une illusion de

¹⁰ Frezal, S., 2016, Risk and statistics: a regulation fallacy, *Working paper PARI*, 35p.

déterminisme, biaisant son raisonnement et altérant sa capacité de décision rationnelle.

- (3) Qui plus est, modèle de risque et risque de modèle deviennent alors indiscernables : il n'est pas possible d'identifier *ex post* de responsabilité (« accountability ») quant à la qualité du modèle. Réciproquement, il permet également de justifier *ex post* n'importe quelle décision, quelle que soit l'issue observée, formalisant la « faute à pas de chance », et déresponsabilise donc également les décideurs-utilisateurs du modèle, leur offrant une protection rhétorique contre la critique.
- (4) Enfin, il n'est alors pas possible de calibrer une fonction de distribution *ex ante* sans arbitraire. Ceci sape l'ambition de tout modèle de risque de permettre d'objectiver les décisions.

Dans un second temps, j'applique ce cadre conceptuel à Solvabilité 2. Ceci permet (i) d'illustrer empiriquement les résultats obtenus et (ii) de proposer un mécanisme général pour un ensemble de dysfonctionnements observés lors de la mise en place de cette réglementation, les plaçant au sein d'un cadre qui met en évidence leur inéluctabilité, ou du moins leur propension à favoriser les phénomènes suivants:

- représentation mentale déterministe faisant paradoxalement disparaître le risque du cadran d'analyse des décideurs,
- déresponsabilisation des acteurs,
- incapacité à définir une interprétation univoque des indicateurs de risque, et partant à les calibrer.

3.2.2 Chapitre 4 : Coût et tarification des risques : de Solvabilité 2 à la question de l'équité des prix des assurances

« *Fairness in Uncertainty: some Limits and Misinterpretations of Actuarial Fairness* »¹¹

Si le débat sur la légitimité du prix d'une assurance existe depuis plusieurs siècles (cf. supra), il est rendu particulièrement aigu par l'émergence du big data qui permet aux assureurs de différencier de plus en plus finement leur tarif entre assurés, soit pour des raisons techniques, tel assuré lui apparaissant plus risqué que tel autre, soit pour des raisons commerciales, tel assuré étant susceptible de présenter une élasticité-prix plus faible que tel autre. De part et d'autre de l'Atlantique, ces évolutions technologiques et les adaptations de régulation induites ont conduit à revivifier le débat sur la légitimité d'une différenciation ou d'une non-différenciation tarifaire.

Dans ce contexte, cet essai rappelle le cadrage théorique développé dans l'essai précédent afin de le déployer sur la question de la régulation tarifaire en assurance. De fait, lors de la souscription d'un contrat d'assurance, les deux acteurs que sont l'assureur et l'assuré sont placés dans des situations différentes. D'un côté, l'assureur, qui souscrit un grand nombre de contrats, est placé en situation plurielle : les grandeurs statistiques sont pertinentes pour déterminer ses orientations tarifaires. De l'autre, l'assuré, confronté à une potentialité d'occurrence de sinistre faible, est placé en situation aléa – à l'instar d'un assureur confronté à son propre risque de faillite, comme dans le cas d'application de l'essai précédent.

Dès lors, ses enseignements généraux sur les conséquences de l'exploitation d'outils statistiques en situation d'aléa peuvent être exploités, au premier rang desquels les conséquences (1) et (4) :

¹¹ Frezal, S. et Barry, L., 2019, *Fairness in Uncertainty: some Limits and Misinterpretations of Actuarial Fairness*, *Journal of Business Ethics* (2019) <https://doi.org/10.1007/s10551-019-04171-2>.

- du point de vue de l'assuré, il n'est pas pertinent de fonder une décision (de tarification, de souscription à un prix donné) sur une grandeur telle qu'une prime pure ;
- et il n'est pas possible de calibrer la prime pure d'un assuré sans arbitraire – en d'autres termes, il n'est pas possible d'individualiser objectivement une prime pure.

Cette perspective conduit ainsi à délégitimer la prime pure comme point de référence pertinent, du point de vue de l'assuré, pour apprécier le bien fondé d'un tarif d'assurance. Dès lors, elle déplace le débat éthique et réglementaire en l'éloignant des considérations de justesse et de justice actuarielles qui en fondent actuellement largement la problématique pour le repositionner sur des enjeux d'efficacité économique et sociale.

3.3 Partie III : Retour sur les facteurs d'institutionnalisation de l'approche statistique promue par Solvabilité 2

3.3.1 Chapitre 5 : La réussite d'un amalgame problématique

« *Instituer l'incobérence. Aléa et hétérogénéité au sein du secteur assurantiel* »¹²

Comment expliquer que les outils statistiques – et l'amalgame entre aléa et hétérogénéité qui les sous-tend – aient percolé au sein du secteur des assurances et s'y soient largement diffusés ? Nous mettons en évidence une première phase d'appropriation, correspondant à la constitution de conglomérats financiers, dans les années 90, lorsque de grands groupes d'assurance mènent des opérations de fusion-acquisition sur d'autres acteurs du secteur financier (e.g. Axa acquérant DLJ, une société de gestion d'actifs, ou ensuite Allianz rachetant la banque Dresdner). A cette occasion, ces grands

¹² François, P., et Frezal, S., 2018, *Instituer l'incobérence : aléa et hétérogénéité au sein du secteur assurantiel*, *Sociologie du travail*, Vol. 60 – n° 1.

groupes se confrontent aux – et endossent les – mesures statistiques de risque (telles que les VaR), notamment comme outils privilégiés de communication aux investisseurs. L'usage de ces grandeurs s'étend d'autant plus que ces groupes multi-activités et multi-pays se rapprochent eux-mêmes de structures de type holding d'investissement. Dans la continuité de ces premiers usages, et capitalisant sur des ressources acculturées à cet amalgame – de jeunes actuaires dont la formation évolue et qui sont désormais formés aux mathématiques financières –, les principaux groupes européens d'assurance promeuvent des nouvelles normes de communication aux investisseurs – au premier rang desquelles la MCEV, qui place l'espérance au cœur de la valorisation des portefeuilles d'assurance. Solvabilité 2 reprend à son compte ces concepts et, les imposant par sa dimension réglementaire à l'ensemble des acteurs, assure leur généralisation à l'intégralité du secteur en Europe.

Dans une seconde partie de cet essai, nous tentons d'éclairer la posture qui accompagne le recours à ces outils au sein de différents acteurs de la profession. Il apparaît alors deux catégories d'approches. Certains font acte de foi, endossant totalement ces outils, convaincus de leur pertinence pour analyser et décider face à l'incertitude des marchés financiers et de l'activité d'assurance. D'autres entretiennent une distance plus réflexive et justifient leur utilité et leur utilisation non tant par leur capacité effective à capter les risques que par leurs vertus pratiques. On retrouve alors dans celles-ci les vertus qui ont fait le succès des statistiques au cours des siècles et, plus récemment, dans les autres secteurs de la finance, au premier rang desquelles les capacités de justification (éteindre la critique) et de commensuration (« *benchmark* » permettant de clore une divergence d'appréciation).

3.3.2 Chapitre 6 : Le basculement vers une logique marchande

« de quoi Solvabilité 2 est-il le nom ? »¹³

Cet essai propose, en premier lieu, une présentation synthétique des principes comptables et prudentiels de Solvabilité 1 et de Solvabilité 2, ainsi que leur mise en regard.

Par ailleurs, et surtout, j'y montre que les choix techniques incorporés dans Solvabilité 2 reflètent des priorités différentes de celles fixées dans ses considérants. En effet, la protection des assurés apparaît, *de facto*, secondaire par rapport à l'objectif d'information aux investisseurs. Par exemple,

- Le *best estimate*, qui constitue la masse des provisions, ne correspond pas à un montant garanti tangible pour l'assuré, mais permet de renseigner les investisseurs sur la façon dont les actifs seront répartis entre la communauté des assurés et eux.
- Une faillite ne correspond pas au franchissement du seuil d'incapacité de tenue des engagements envers les assurés, mais à la percussioin du *strike* (nul) de l'option de vente sur fonds propres économiques qu'est l'action détenue par les investisseurs.
- Le « risque » mesuré ne correspond pas tant à une perspective de pertes qui mettrait en danger la compagnie et ses *stakeholders*, mais à un indicateur de sensibilité qui peut croître lorsque la santé de la compagnie s'améliore (dominance simple), apparentant le ratio de solvabilité à un ratio de Sharpe, outil d'appréciation relative d'investissements, plus qu'à un ratio prudentiel qui aurait été une fonction croissante de la capacité de la compagnie à tenir ses engagements vis-à-vis des assurés.

¹³ Frezal, S., 2015, De quoi Solvabilité 2 est-il le nom ?, *Working paper PARI*, 56p., devrait paraître comme chapitre au sein d'un ouvrage collectif aux Presses de Sciences Po.

Ainsi, Solvabilité 2 semble avoir eu pour critère cardinal l'amélioration de l'information financière : la logique civique de protection des assurés a, *in fine*, été englobée dans une logique marchande de développement d'un marché européen des passifs d'assurance par la mise en place d'outils de commensuration de ces « actifs » et de leurs rendements.

Solvabilité 2 étant entrée en vigueur en 2016, il est encore trop tôt pour savoir si cet objectif sera atteint. En revanche, les exercices préparatoires, débutés à la fin des années 2000, ont permis à un certain nombre d'acteurs de s'acculturer à ces outils. La deuxième partie de ce chapitre esquisse les premières pistes d'évolution vers lesquelles l'introduction de ces pratiques semble orienter, à ce stade au niveau du pilotage interne des compagnies. A la lumière de ces éléments, des effets pervers apparaissent, tels qu'un alignement des mentalités susceptible d'accroître le risque systémique, ou une disparition du risque du cadran d'analyse susceptible de diminuer la conscience des périls. Des études ultérieures devront toutefois confirmer ou infirmer ces observations, ainsi que leur imputation à la mise en application de Solvabilité 2.

Partie I : Les limites de l'approche statistique adoptée par
Solvabilité 2

Chapitre 1 : Le problème de la fiabilité des estimateurs

Solvency II Is Not Risk-Based—Could It be ? Evidence from Non-Life Calibrations

Introduction

From a theoretical point of view, the principle of basing capital requirements on risk measures rather than on volume indicators has been debated, with some authors considering it effective (e.g., Cummins & Phillips, 2009; Eling & Holz Müller, 2008; Weber & Darbellay, 2008) and others taking a more circumspect stance (e.g., Repulo & Martinez-Miera, 2014 regarding banks; Author et al., 2016 regarding insurance). This article does not seek to take a direct position in the debate – we begin from the established fact that, in insurance, the Solvency II Directive (EPC, 2009), which redesigned prudential regulation for insurance in Europe and which is spreading in numerous countries over the world, is meant to be *risk-based*:

The starting point for the adequacy of the quantitative requirements in the insurance sector is the Solvency Capital Requirement (SCR). [...] The Solvency Capital Requirement standard formula is intended to reflect the risk profile of most insurance and reinsurance undertakings. (Art. 26)

Where Solvency I capital requirements were considered too unsophisticated (for example, only distinguishing two sets of lines of business in non-life insurance, weighting the riskiest by a fixed factor of 1.5, and then simply adding their contributions to determine the capital requirements), Solvency II discerns 12 lines of business for non-life insurance, for each of which it calibrates shocks including more than two significant figures, and then aggregates these, together with other risks, taking into account a dependence framework to determine the capital requirements (see Pfeifer and Strassburger, 2008, for a presentation and critical analysis of this step).

This refinement of capital requirements characterization is supposed to shed light on two fields. That of regulatory action (licensing and withdrawal), as these measures directly impact the denominator in the solvency ratio. And that of company steering, since these measures directly impact the denominator in

the profitability ratio: for insurers and investors seeking to arbitrate between different lines of business, their risk-adjusted profitability should be reflected by these measures. Our goal is to determine whether these refined calibrations are sufficiently robust for the various economic agents to base their operational decisions on them. Indeed, this specific ambition to be risk-based faces three kinds of difficulties:

Scientific. The risk models have yet to prove their robustness. In banking, the FSA observed that banks' internal models diverged by a factor of 6 for a representative portfolio of assets (Harrison & Samuel, 2011). Danielsson (2002, 2008) has also demonstrated that, by retaining a very common range of historical depths and a limited and standard set of underlying mathematical models, we observe up to a two-fold difference between the different possible estimations of a daily VaR at 99% for a liquid vanilla security. In insurance, within the framework of Solvency II, this model risk could be much higher than in the example above. Indeed, (1) the targeted VaR and the data used are not daily, but annual: the lower frequency means that it is even more necessary to extrapolate on the basis of a limited sample size, which degrades the robustness of the statistics that can be estimated, and (2) certain data, such as those derived from claims and related reserving, must be systematically preprocessed before being used, which increases recourse to expert judgments and thus expands the range of non-illegitimate final results.

Interpretive. Even if the mathematical definition of the risk measure is univocal (an annual VaR at 99.5%), the physical meaning associated with this quantity, as well as the associated calibration methods, are more ambiguous. As such, we do not know whether this probability should be interpreted as "each year, one company in 200 goes bankrupt" or "every two centuries, the entire market goes bankrupt". These two formulations represent polar cases, which, in the first instance, would lead to quantifying only specific risk and, in the second,

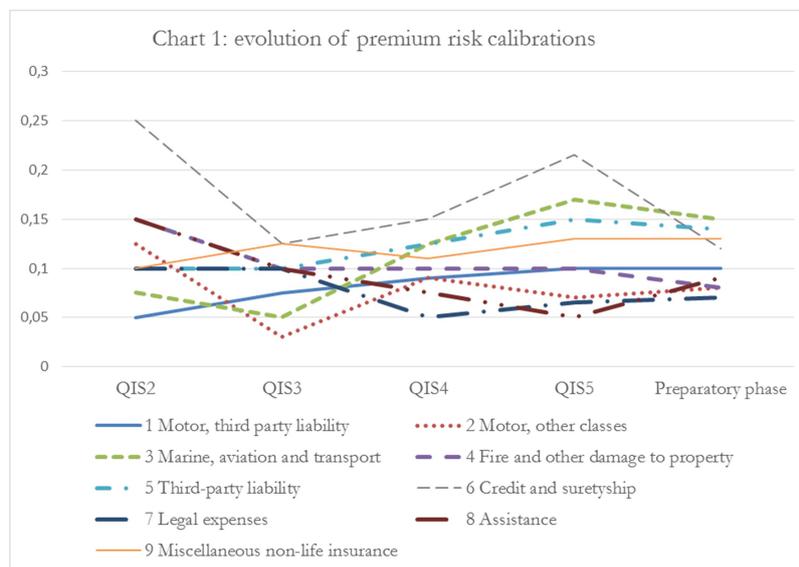
only systemic risk. They constitute frames of mind, as neither the different accounting periods nor the different companies are independent, but privileging one interpretation or another leads to different quantifications – in the one case, concerned with differentiating the behaviors of the different companies, and in the other with preventing the impact of a systemic crisis. Leroy and Planchet (2010) thus interpret a period of upward adjustment of calibrations as the mark of a movement to fight against systemic risk¹⁴.

Political. While calibrations are considered to be technical (simple econometric measures), they also contain political choices. The most flagrant example (which is criticized) is that there is no postulated risk for sovereign debts in the quantitative framework (Pillar one) of Solvency II. For example, Greek public debt was, and continues to be, considered to be risk free: its annual VaR at 99.5% is considered zero (which is understandable, as public authorities have legitimate concerns about the consequences of prudential regulation on the financing of the economy and the stability of the markets, but which does not fit with the claim of being risk-based).

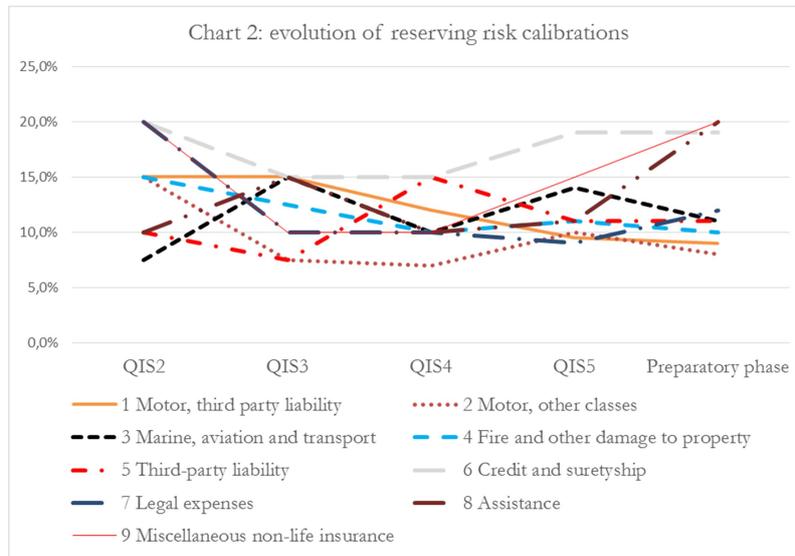
These three points impose serious limitations to Solvency II's ambition to offer a framework "based on risks". Awareness of some deficiencies of the quantitative measures exist: the theoretical limits of the VaR aggregation have been explored (Embrechts et al., 2013, Pfeifer, 2013), some sensitivity measures have been produced (El Karoui et al., 2015) and regulators facing a diversity of modelling among companies empirically observe that the quantitative measures are not that sound. Nevertheless, these mainly qualitative

¹⁴ We can formulate this concrete interpretive ambiguity of VaR and the associated calibration differently by taking for example the probability of the French state going bankrupt. If we consider it impossible that the French state will go bankrupt this year, but not exclude that it might do so in the next 200 years, should we consider the annual VaR at 99.5% to be zero (the organization will not be affected by this risk next year) or non-zero (the market may be affected by such an event in the next 200 years)? While the mathematical definition appears to be univocal, we find ourselves at an impasse when interpreting the quantity in question, and hence when determining how to measure said quantity.

critiques are usually dismissed with thoughts such as “*don’t throw out the baby with the bathwater*”, “*it’s a proxy, a good proxy*”, “*the order of magnitude is right*”, or “*it’s not perfect, but it’s better than nothing*”¹⁵. However, the high volatility observed in the calibrations leads one to wonder whether these qualitative responses are well-founded. For example, in 2006, the European regulator (CEIOPS) considered third-party liability to be riskier than transport insurance, then in 2007 decided the contrary (twice less risky), then in 2008 returned to its initial hierarchy (1.5 more risky), reversing the order once more in 2010...then finally considering them equivalent in 2014. More generally, charts 1 and 2 present, for premium and reserving risks (which together represent 40% of the capital requirement for non-life insurance companies; cf. ACPR, 2016, p.16), the estimated level of risk for each line of business (LoB) over time, as carried out by the CEIOPS in the context of preparatory work for Solvency II. Each intersection of two lines marks a change in opinion regarding the relative risk of two LoBs.



15 Another regular position to dismiss criticisms is that it would not matter whether things are (mathematically) correct or not, as long as all firms and countries do the same (wrong) things. However, should unicity be the core concern, it would become very hard to justify the switch from Solvency 1, which already set up a capital requirement framework unified among firms and countries (which was much more simple).



In order to characterize this volatility with a synthetic indicator, we may measure the proportion of LoB couples whose hierarchies remain stable throughout the estimations. Doing so, we find that in only one-third of the cases (31% for both premium risk and reserving risk) the assessment of one LoB as riskier than another remains stable over the course of the studies. As practitioners review the estimations in such a way, we doubt the ability of this technology to effectively capture risk.

Danielsson (op. cit.) in the field of banking capital requirements, Kamega & Planchet (2013) and El Karoui et al. (op.cit.) in regard to the available economic capital of insurers¹⁶, have tested the impact of changing the set of inputs, thus exploring the quantification of error margins. Albeit it provides valuable insights, it is exposed to the following criticism: *“I do believe that my/this choice is more appropriate than yours/this other choice. As such, you are not measuring an intrinsic margin of error but the error made by these who do not measure in an appropriate way”*. I developed an alternative methodology, based on the observation and comparison of different official calibration sets devised by regulators and

¹⁶ Through life reserving and thus, by complementarity, the available own funds.

industry actors, which is not exposed to such objection. I compare the dispersion of panel risk measures along two dimensions: different assessments of a given risk provided by an identified actor (e.g. the regulator), and the assessment of different risks provided by this actor. This allows measuring the relative amplitudes of noise and signal where, in keeping with the expression *signal-noise ratio* used by Sims (2003) and Alesina & Tabellini (2007), signal corresponds to the characteristic amplitude of information that a measure seeks to register and noise to the characteristic magnitude of disturbances that muddle this information.

The relative assessment of noise and signal should enable us to dismiss (or not) a measure as providing operationally useful information. Let us take an example: suppose we use an altimeter to fly a plane. If, due to the characteristics of the plane at the beginning of the 20th century, it cannot fly higher than 1000m and the altimeter is only accurate to within 10,000m, then the amplitude of the signal we seek is low regarding the interference noise, and the altimeter has no operational utility (cf. figure 1, case B). Inversely, if the plane can fly up to 10,000m and the altimeter is accurate to within 1000m, it is useful in certain circumstances (not for landing, but at least for avoiding a mountain; cf. figure 1, case A). In case A, where noise is low compared to signal, no systematic conclusion can be reached: the width of the margin of error (e.g. the volatility of the measure) is an issue or not depending on the intended operational use. But in case B, where noise is high compared to signal, the measure is useless whatever the precision required by the intended use may be. In other words, considering a signal-noise ratio allows, when it is poor, to reach a general conclusion disqualifying the steering tool as inappropriate whatever its operational use may be, without requiring to consider the specific decision making context.

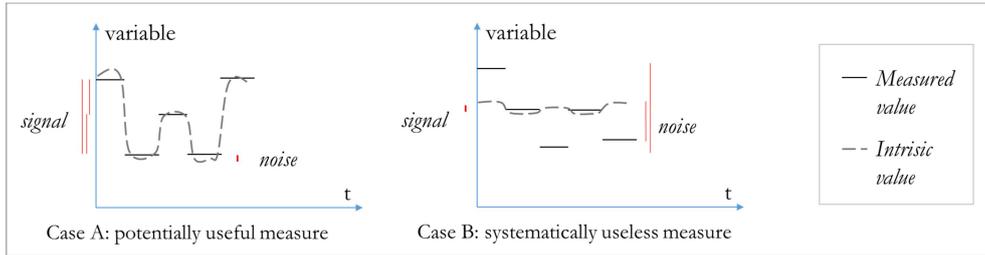


Figure 1: signal-noise visualization of a measure's operational pertinence

Furthermore, considering the relative magnitude of noise and signal has the advantage of being agnostic as regards the claim of objectivity of a specific measure: it measure the internal credibility of the measurement system. Reciprocally, our methodology has the disadvantage of not being able to prove the validity of a measurement system: it cannot confirm the operational pertinence of calibrations, but only enables us to deny it, when applicable.

In the first section of this article, we will apply this methodology to determine whether these observations should lead us to consider that the calibrations do not provide operationally useful information or whether, on the contrary, they may be “better than nothing”. In order to do this, we will first precise the methodology employed to evaluate a signal-noise ratio; we will then specify the data used; and finally, we will present the results and their interpretation.

In the second section, we will go through the typology of sources of the margin of error to analyze the different possible causes of this situation. Relying on the informations and data provided by the regulator to support and put into perspective its calibrations, we will try to determine the significance of each: political “pollution” of a technical measure, or intrinsically technical limits, either due to the methodology or to the underlying data. This will allow us to propose prudential design improvements.

1. Are these risk measures “better then nothing”?

1.1 Methodology

To assess the dispersion of measures relating to noise and to signal respectively, we compare (see figure 2):

- a *noise* indicator that reflects the dispersion of alternative calibrations of the same quantity (this is characteristic of the difference between evaluations of a *given risk* assessed through *different approaches*)
- to a *raw signal* indicator that captures the dispersion of calibrations for different quantities (this is characteristic of the difference between evaluations of *different risks* by a *given approach*).

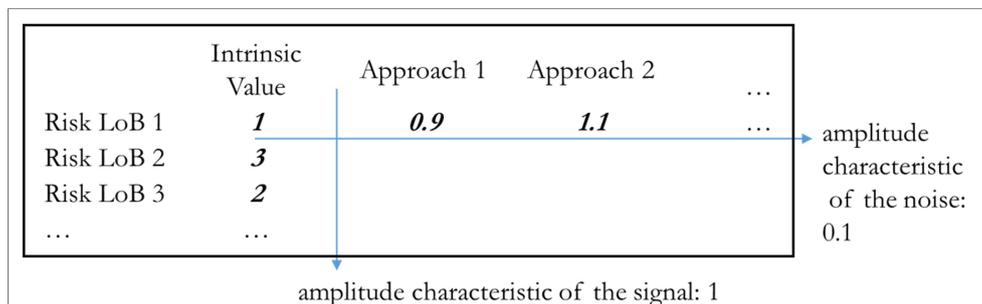


Figure 2: Illustration of signal and noise

These dispersions are measured for a set of risks and a set of approaches with the aid of two common and intuitive indicators:

- the standard deviation
- and for the “extreme” dispersion, the ratio $\frac{\max - \min}{\min}$.

This second indicator can only be positive, and if it is greater than 1, this indicates there is more than a two-fold difference between the lowest and highest estimates.

Idealistic world

For the sake of clarity, let us assume first that we have access to the intrinsic value of the parameters. Then, we can determine the amplitude of the signal, the amplitude of the noise, and straightforwardly assess the signal-noise ratio by dividing the first by the second.

This allows to get the intuition of the ratio. It ranges from 0 to infinity. When high (say 10, 100 or 1000), the precision of the measure and its discriminating capacity are good (precision of 10%, 1% or 0.1%). When equal to 1, it means that the magnitude of the noise is comparable to this of the signal: then, half of the apparent discriminative power provided by the measure is in fact nothing but a random pollution. When closer to zero, the signal is very weak, drown in the noise.

However, in practice, we do not have access to the intrinsic value: we just have several sets of measures, each approach incorporating some noise. We shall then take into account the fact that the raw signal indicator, corresponding to the measure of dispersion within a given approach, is itself affected by noise. Otherwise, the “idealistic world” measure would overestimate the signal and lead to overconfidence into the measurement system.

Real World

To refine the analysis, we can write the calibration of the risk $R_{i,j}$ associated with the LoB i in the set of calibrations j in the form

$$R_{i,j} = V_i + n_{i,j}$$

Where V_i represents the value of the risk parameter sought for the LoB i and $n_{i,j}$ represents the noise associated with the LoB i for the set of measures j .

This noise $n_{i,j}$ may comprise:

- LoB-specific components. For example, when one LoB is strategic for public authorities looking to develop this market (its calibration will be revised downward). Or a systematic bias of underestimation of a rare (and therefore unseen) jurisprudential change affecting the reserves of a specific LoB. Etc.
- calibration set-specific components. For example, the use of a new methodology that takes into account the possibility of fat-tailed laws that would lead to augmenting all calibrations. Or a decision of public authorities to reinforce or mitigate the pressure over the whole insurance sector, given the macroeconomic context. Etc.
- a residue, corresponding, for example, to when, after a particular event, judgment about a given line of business is revised, before new information refutes said judgment. Etc.

As such, it can be modeled from three components. If they interact additively, for example, in the form:

$$n_{i,j} = L\varepsilon_i + Q\sigma_j + B\tau_{i,j}$$

where ε_i , σ_j and $\tau_{i,j}$ are normalized white noises respectively reflecting the random component:

- for ε , related to the LoB,
- for σ , related to the calibration set
- and for τ , non-specific,

and L , Q , and B represent the amplitude of each of these components.

Then, three polar cases are to be considered.

First, if, we assume $B \gg L$ and Q , the error matrix is isotropic: we can rewrite $n_{i,j}$ in the form $n_{i,j} \cong N \mathbf{n}_{i,j}$, where $\mathbf{n}_{i,j}$ is a normalized white noise. To get the magnitude of the signal, we should then correct the dispersion corresponding to the raw signal indicator with the dispersion induced by the noise of characteristic amplitude N , as measured by the noise indicator. This leads to consider, instead of the simple signal-noise ratio, the *adjusted signal-noise ratio* (SNR_a) as follows:

$$SNR_a = \frac{\text{raw signal} - \text{noise}}{\text{noise}}$$

Thus, our adjusted indicator is such that if all the cells in the table (figure 2) are uniformly filled by a random generator, the raw signal and noise indicators being asymptotically identical, then the expected value of the signal-noise ratio would be zero, which reflects the fact that the perceived raw signal would just be noise without information¹⁷.

Second, in the case where $L \gg B$ and Q , which is to say that “the error comes from the LoB”, the noise measured for each LoB between different approaches is low in comparison to the noise between LoB mistakenly incorporated into the information in the raw signal measurement: as such, the above adjustment is insufficient. Therefore, by considering the *adjusted signal-noise ratio* proposed above, we overestimate the information provided relative to the noise and we are conservative regarding the credibility of the system: should it appear poor, the situation is even worse.

¹⁷ whereas if we retained a raw ratio (raw signal to noise) in this case, we would obtain a misleading result of 1, implying that there is a signal of the same amplitude as the noise.

Third, if we take $Q \gg L$ and B , which is to say that the main source of noise comes from the different approaches and affects all classes homogeneously, then we would have:

$$R_{i,j} \cong V_i + Q\sigma_j$$

In this case, our noise indicator corresponds to \bar{Q} , and the entire dispersion for a column corresponds to information. Thus, our raw signal indicator, which corresponds to $\overline{dispersion}_i(R_{i,j}) = \overline{dispersion}(V_i)$, effectively captures the amplitude of the signal. Thus, the *unadjusted* raw signal-noise ratio, by construction equal to $1 + SNR_a$, is to be considered as the significant signal noise ratio, just as it was the case in the idealistic world.

- (0) To summarize: In general, considering the raw signal-noise ratio is, depending on the structure of the noise, either appropriate or conservative.
- (1) If the noise comes from the approach and is homogeneous among the different LoBs, then it is appropriate.
- (2) Otherwise, it is conservative. And then, considering the adjusted signal-noise ratio (SNR_a) is either appropriate (isotropic error matrix) or conservative (if the noise comes from the LoB and is homogeneous among approaches).

Given the data, we do not need to stay in the general case (0): as appears in graphics 1 and 2, the evolution of the calibration for the different LoBs are not parallel. Therefore, we are not in case 1. This suggests that we consider rather the adjusted signal noise ratio, which will provide either appropriate or conservative results (case 2).

1.2 Data

What different kinds of approach can be used to measure a given risk? We will look separately at two categories of approach: one being those approaches developed by the European regulator to calibrate the *standard formula*, the other being those developed by a large insurer for its internal model.

1.2.1 QIS

Solvency II was subjected to a series of quantitative impact studies (QIS) designed to measure the capacity of organizations to perform calculations, to assess the impact of this regulation's implementation on their need for financing, and to refine calibrations. Four of these studies (QIS 2, 3, 4, and 5) proposed a calibration of the 1-year 99.5% VaR calibrations for a series of risks, and the final implementing measures led to a final set of calibrations. As such, we have five sets of risk measures at our disposal, each representing a "possible approach" determined by the CEIOPS or the European Commission.

The conformism (generated, e.g., by market standards¹⁸) biases the analysis by restricting the range of calibrations in relation to the set of those that are "scientifically legitimate". The observed dispersion between approaches is thus underestimated, generating an artificial convergence of results and thus favoring an image of the system as possessing internal credibility. To limit this bias, we do not consider calibrations of market risks, where the use of VaR has long been widespread and thus methods have had more occasion to be standardized. Among underwriting risks, we focus on premium and reserving risks – those for which we have a significant number of standardized stresses,

¹⁸ The directive explicitly emphasizes the necessity that, in terms of calibration, the "specification is in line with generally accepted market practice" (EPC, 2009, Art. 122.4). To give an example in terms of implementation, the *QIS 3 calibration paper* (CEIOPS, 2007b, p. 12) explains that for the purpose of estimating VaR at 99.5% for non-life premium risk, a value of 3♦ is retained, "assuming a lognormal distribution of the underlying risk". This hypothesis applies to every LoB considered and thus tends to guide and confine the choices of laws selected from all of the possible extrapolations.

which allows us to conduct our comparative study. We thus focus on the risks of the 9 LoBs of direct insurance and proportional reinsurance.

1.2.2 An internal model

Those insurance companies that prefer to use an internal model may opt to do so. We have at our disposal calibrations measured internally by an insurance group, at different dates and for each of the risks considered. The group in question is the French sub-group of one of the four global European insurance groups, which developed an internal model and is present in each of the LoB.

We have usable data for reserving risk¹⁹:

- on the five main LoBs (motor, third-party liability; motor, other classes; marine, aviation and transport; damages; and third-party liability),
- over six years, from the end of 2009 to the end of 2014. The insurer had developed a model over several years preceding 2009. The methodologies thus had time to crystallize before the data range we have here, which is likely to create an artificial stability. We therefore certainly underestimate the margin of error more significantly than with QIS data.

1.3 Results

1.3.1 QIS

Figures A3 and A4 (Appendix) present the set of calibrations. Figure 3 synthesizes these results, presenting the averages of the noise indicators (average across LoBs of the dispersion between approaches) and signal indicators (average across QIS of the dispersion between LoBs).

¹⁹ The methodology for premium risk differs from that of the standard formula in that it focuses directly on the extreme quantile and not on the difference between the extreme quantile and the best estimate. The interpretation of the risk measure thus does not correspond to a dispersion, and comparisons between classes are no longer directly valid, as each one has a different expected value. We do not have access to these expected values to provide us with comparable “volatility” indicators.

	st. dev.		max/min	
	Premium	Reserving	Premium	Reserving
Raw signal	4.10%	3.70%	2.7	1.3
Noise	3.20%	3.40%	1.4	0.9
SNRa	28%	12%	99%	47%

Figure 3: average indicators of raw signal, noise, and adjusted signal-noise ratio (QIS)

The following elements appear to stand out:

- On the basis of the first dispersion indicator (standard deviation), the raw signal essentially captures noise. For the premium risk measure, the adjusted signal to noise amplitude is less than 30%. For the reserving risk measure, this signal-noise ratio becomes utterly negligible.
- The results obtained from the second dispersion indicator (max-min/min ratio), though not as pronounced, tend to confirm this first observation. Thus, for premium risk, the signal-noise ratio is 1 and for reserving risk it is 0.5.

It appears difficult, then, to consider that Solvency II's proposed risk measures offer a significantly higher dispersion than that of an outside interference (be it technical or political).

1.3.2 Internal model

For non-technical reasons, we could expect the calibrations resulting from an internal model to be significantly more stable than those resulting from the standard formula's different tests: first, because of the desired stability of a methodological framework that is to be submitted to a supervisor's approval; second, due to their relative immunization to the fluctuations of political orientations to which European text drafts are subjected; and third, because the data we have been able to use originate several years after the establishment of an internal model and the production of associated figures by the entity in

question, allowing for an initial crystallization. As such, the perceived margin of error could be very low for these data.

Nevertheless, reproducing the same studies leads to results that, even if they are indeed slightly better, still show internal noise to remain considerable:

	st. dev.	max-min/min
Raw signal	10.9%	1.8
Noise	6.5%	0.9
SRNa	68%	106%

Figure 4: average dispersion of signal and noise using an internal model (reserving risk, 2009-2014)

We observe that:

- on average, the estimation of risk associated with a given class changes by a factor of 2 over the period (max-min/min=0.9),
- the amplitude of the internal signal is less than or comparable to that of the internal noise.

This appears to suggest that these measures do not provide reliable operational information.²⁰

1.4 Interpretation: is the “noise” truly noise?

As we mentioned in the introduction, several sources of diversity are likely to make calibrations vary from one exercise to another:

- a change in political trade-off, related for example to pressure exerted by an industry, or to a member state’s desire to not put “its” actors at a disadvantage, or to the fear of an overly heightened calibration’s consequences on the market, etc. These political assessments were

²⁰ The same holds from an ordinal point of view, as the stability indicator for the hierarchy of the five classes over the 6-year period considered is 40%. This means that a judgment on the relative risk serving as the basis for a decision will have been inversed in most cases.

encouraged, potentially, by the fact that the deterioration of the financial market at the time added a strain to companies' capital positions and their ability to raise funds should the new regulation require it.

- a technical evolution (or a change in the interpretation of the risk measure), either (i) related to the data themselves, for example on the occasion of a refinement of the data, an increase in historical depth or, on the contrary, a change in products that renders a past calibration obsolete, or (ii) related to expert judgments, for example due to the use of a new kind of model (e.g., a change in the underlying distribution function), or simply a change of personnel responsible for the calculation (to illustrate that point, we can notice that a manager who knows his team can generally tell which team member has conducted a statistical analysis by observing which “outlying” points have been excluded from the analysis).

The reason for this volatility, be it technical or political, is of little importance. If it is political, this obviously distorts the analysis' purely technical credibility. If the reason is technical, it seems unjustifiable to consider that the insurers' data and the field of actuarial science made a qualitative jump sometime between 2006 (the date of QIS 2) and 2010 (the date of QIS 5) and that, as opposed to 2006, 2010 or 2014 would represent the year when the actuarial science and the available data reached their full realization.

Two arguments may yet explain a dispersion that might be wrongly interpreted as noise:

- The first argument is that risks change and thus it is legitimate that calibrations adapt to them. Our analysis does not permit us to rule out this argument. However, (1) the period of revision for the calibrations

observed (which is annual for the first few QIS) is far smaller than the duration of the liabilities under consideration and the associated risks (approximately 5 years on average), and (2) if prior changes to the calibrations did in fact reflect the changes in the underlying risk, their crystallization would generate a gap between the calibration and the underlying risk of a comparable order of magnitude in the future. And there are no plans to revise the directive's calibrations on a regular basis.

- The second argument is that certain of the calibrations from the first few QIS were based on data from some national markets within the European Union (CEIOPS, 2010b, p.189), and thus the expansion of the database could explain the variations. Once again, the quantitative pertinence of this argument cannot be discounted by our analysis. However, if this were to be the source of the variations, it would mean that the heterogeneity of the underlying risks among national markets is such that imposing uniform measures generates a gap between the underlying risk and its measure at an order of magnitude comparable to that of the “noise” measured. Thus, it is effectively noise that we find when we attempt to determine the risk profile of a given company, which is the objective and claim of Solvency II.

As such, in the context of the Solvency II Directive, whose calibrations are set to be fixed and homogeneous, the arguments above do not call into question the consideration of the dispersion between the different QIS (and with the Level 2 implementing measures) as a matter of noise – of an undue discrepancy between a company's capital requirement and its risk²¹. This leads

²¹ As concerns internal model, the sources of variations both over time and in regard to the standard model may be of a political or a technical nature, or both. If they are political in nature, then it is a different political nature than that of the standard formula: here, this would be the company's internal issues (e.g., supporting a strategic choice or steering an internal message towards the holding company or external actors such as rating agencies). If such

to a validation of the interpretation of instability as an unreliability. Is it possible to identify the sources of this noise and to adapt the design of prudential regulation?

2. The causes

In this second section, we study three explanations that may account for the observed instability of the calibrations: the possible disruption of technical calibrations by political issues, the importance of the idiosyncratic factor, and a potential epistemological barrier. To this end, we will systematically rely on the data and the methodological elements provided by the CEIOPS to support its proposed QIS calibrations.

2.1 Political economy?

A possible explanation to the volatility of the calibration is the political interference. Can we validate such hypothesis that political choices are significant sources of variation in calibrations?

2.1.1 Data used

The CEIOPS has occasionally accompanied QIS calibrations with supporting evidence:

- No evidence was provided for QIS 2 (CEIOPS, 2006).

dispersion is related to expert choices, this destroys the very purpose of a scientifically established risk-based perspective; if it is related to a change in the data, this contradicts the notion that a standard formula (stable over time) could be pertinent. Moreover, in this case, it appears that the period of obsolescence for the information provided by these calibrations is lower than the duration of the portfolio, which raises doubts as to the pertinence of their use in steering by companies.

Once more, the reasons for the dispersion of parameters deemed to be technically pertinent is of no matter; the dispersion of observations is to be classified as noise and leads to a disqualification of these calibrations as reliable tools for steering and regulation.

- Almost no evidence was provided for QIS 3 (CEIOPS, 2007b). However, some information suggests that this new set of calibrations fits in a global context of expectations of stress reduction.²²
- Supporting evidence is nearly inexistent for QIS 4, with the CEIOPS indicating, for example, that “*attention was paid to the industry's QIS 3 feedback regarding the calibration of the SCR formula, but being globally pleased with the QIS 3 calibration, CEIOPS decided not to substantially challenge the QIS 3 calibration*” (CEIOPS, 2008).
- However, in 2009 (an abortive calibration considered at the time as the CEIOPS’ final proposition to the European Commission, which did not give rise to an impact study) and in 2010 (QIS 5, effectively the CEIOPS’ final recommendations to the European Commission), the CEIOPS provided information about the calibrations and the methodology employed (CEIOPS, 2009 & 2010b).

For each of these two exercises,

- (i) the CEIOPS applied four estimation methods for premium risk and six for reserving risk to the data at its disposal (a few countries in 2009, and a larger perimeter in 2010). All of the methods used were based on the same assumption of underlying law and differed by the number of degrees of freedom retained.²³
- (ii) Then, risk by risk, the CEIOPS identified which method(s) seemed the best fit(s),

²² For example, between QIS 2 and QIS 3, the closed formula enabling the transformation of the volatility measure into a regulatory risk measure shifted from a formula reflecting a TVaR at 99% to one reflecting a VaR at 99.5% (both based on the same premise of an underlying log-normal distribution), thereby generating a decrease of 5 to 10% in the regulatory calibration for an unchanged volatility.

²³ e.g., (1) the calculation of a unique expected value and a unique volatility, (2) the calculation of a unique expected value and of a volatility per company (to be averaged), (3) the calculation of an expected value per company (to be averaged) and of a unique volatility, etc.

- (iii) and finally chose (a) certain method(s) to determine its proposed calibration of the standard deviation for each of the risks (sometimes the average of the methods retained, sometimes a figure deemed “consistent” with the adopted methods).

2.1.2 Results

The set of methods adopted for the calibration proposal (step iii) regularly deviates from the set of methods considered most pertinent from a technical standpoint (step ii). Focusing on premium risk, LoB by LoB, figure 8 synthesizes the set of methods considered most pertinent and the set of methods adopted for the final calibration: there is no LoB for which these two sets are identical.

Premium risk	CP 2009		QIS 5	
	best fit	retained	best fit	retained
1 Motor, third party liability	3-4	1 - QIS4	4	1-4
2 Motor, other classes	3-4	3 - 1 -QIS 4	2-4	1-2-4
3 Marine, aviation and transport	2-3	4-2	2-4	1-2-4
4 Fire and other damage to property	2	1 - QIS4	none	1-2-3-4
5 Third-party liability	3	?	4	1-4
6 Credit and suretyship	2	?	2-4	1-2
7 Legal expenses	2-4	?	2-4	1-2
8 Assistance	2-4	2 - QIS 4	2	1-2
9 Miscellaneous non-life insurance	2-4	2 - QIS 4	2-4	1-2-4

Figure 8: methods considered optimal and methods adopted for the calibration

The same phenomenon can be observed in regard to reserving risk. For example, the consultation paper (CP, 2009) identifies three “pertinent” estimations of 25%, 23%, and 45% for LoB 9, then concludes without justification that the calibration should be 20%.

Generally speaking, the justification for this deviation (passage from step (ii) to step (iii)) is often very weak, if not non-existent. For example:

“Overall conclusions: Method 4 and 2 provide a good fit. This would imply a factor of 14% on average based on the fitted results. Method 5 does not allow for diversification and in views of the graph above seems to ignore some important observations. A final factor considering method 5 and method 1 has been selected.” (CP2009, reserving risk, LoB 7).

It is possible to estimate the total impact of this divergence between the “technical optimum” and the adopted proposal. Figure 9 compares, for our representative company, the average estimation of risk according to the methods judged to be pertinent to that which results from the final calibration proposal. The proposed calibrations result in a significantly weaker estimate than the one to which the purely technical calibrations would have led.

<i>Weighted average</i>	Best fit	Proposal
QIS 5	15.1%	13.0%
CP 2009	22.8%	16.3%
Reminder of QIS 4		13.9%

Figure 9: LoB volume-weighted averages of standard deviations linked to premium risk (gross of reinsurance), estimated by the CEIOPS

It is clear here that the successive calibrations took into account a significant non-technical dimension. Does this mean that if the calibrations are not reliable as measures of risk, it is due solely to the interference of political arbitrations distorting the math? Would improving the independence of technical authorities be enough to restore technical credibility to this regulation that claims to be risk-based?

2.2 An idiosyncrasy?

An explanation for the dispersion of calibrations observed between different QIS may be the level of idiosyncrasy on the market. Indeed, let us recall that the first couple QIS calibrations were carried out with limited data, sometimes

restrained to one national market. When firms are very heterogeneous, it is normal for a modification of the calibration sample to alter the parameter.

Certain calibration methods led the CEIOPS to estimate a VaR for each of the companies whose data it had obtained. Figures 5 and 6 show, for those of these methods deemed pertinent by the CEIOPS, the first and third quartiles of premium risk (figure 5) and reserving risk (figure 6) estimated for each LoB.

Premium	Method	75th percentile	25th percentile	max/min
2 Motor, other classes	2	18%	8%	2.3
3 Marine, aviation and transport	2	109%	28%	3.9
4 Fire and other damage to property	2	61%	16%	3.8
4 Fire and other damage to property	3	96%	25%	3.8
6 Credit and suretyship	2	124%	40%	3.1
7 Legal expenses	2	27%	11%	2.5
8 Assistance	2	14%	6%	2.3
9 Miscellaneous non-life insurance	2	77%	15%	5.1

Figure 5: dispersion of estimated volatility between the first and third quartile of firms, according to the methods adopted for the final choice and determining one volatility per company (premium risk, QIS 5)

Reserve	Method	75th percentile	25th percentile	max/min
1 Motor, third party liability	1	17%	6%	2.8
1 Motor, third party liability	2	40%	10%	4.0
2 Motor, other classes	1	40%	14%	2.9
3 Marine, aviation and transport	1	63%	32%	2.0
3 Marine, aviation and transport	2	365%	43%	8.5
4 Fire and other damage to property	1	40%	13%	3.1
4 Fire and other damage to property	2	81%	24%	3.4
5 Third-party liability	1	50%	13%	3.8
6 Credit and suretyship	1	81%	29%	2.8
8 Assistance	1	87%	29%	3.0
9 Miscellaneous non-life insurance	1	72%	25%	2.9

Figure 6: dispersion of estimated volatility between the first and third quartile of firms, according to the methods adopted for the final choice and determining one volatility per company (reserving risk, QIS 5)

We see that the estimated risk level differs by a factor of 3.5 on average between the first and the third quartile (it varies by a factor of 2 to 8 depending on LoB and method).

This dispersion of risk measures between companies within the same LoB is much higher than the dispersion between LoBs that Solvency II claims to capture (for premium risk, the class with the highest risk (MAT, 15%) differs from that with the lowest risk (legal expenses, 7%) by a factor of 2; and by a factor of 2.5 for reserving risk). As such, the choice of tacking a standard formula – calibrated by LoB but homogeneous between firms – onto companies (that are so heterogeneous within the same LoB) provides no significant added value in terms of risk measurement compared to a system that treats classes homogeneously.

This suggests that the changes of scope of data between QIS may have had a significant impact on the volatility of the calibrations, without this calling into question the “noisy” character of these developments. More important, it shows that to be risk based would require to generalize the practice of undertaking specific parameters (USP) that aim to capture the risk specific to each company. However, if this were necessary to achieve Solvency II’s objective, would that be enough?

2.3 An epistemological barrier?

Independent of the political interference and the idiosyncrasy to which we have just called attention, the analyses provided by the CEIOPS to support their calibration proposals show that, using a constant dataset, different methods lead to widely dispersed results. Figure 10 highlights the dispersion of the calibrations (ratio between the highest and the lowest) for each LoB according to the method utilized: depending on the LoB and the exercise considered, the gap between estimates varies by a factor of 2 to 60.

<i>ratio max /min</i>	Premium		Reserves	
	CP2009	QIS 5	CP2009	QIS 5
1 Motor, third party liability	2.4	4.2	7	13
2 Motor, other classes	2.2	2.7	5	7
3 Marine, aviation and transport	3.3	3.1	7	7
4 Fire and other damage to property	2.8	1.8	6	11
5 Third-party liability	2.6	1.8	23	22
6 Credit and suretyship	6.0	3.2	5	60
7 Legal expenses	8.0	17.3	12	16
8 Assistance	6.0	5.5	34	26
9 Miscellaneous non-life insurance	48.0	4.1	2	20

Figure 10: ratio between extreme estimates provided by the different methods used

Even when restricted to the only methods adopted by the CEIOPS on a case-by-case basis to determine the final calibrations, the dispersion remains very high – up to a factor of 8 (cf. figure 11) – much higher than the dispersion of calibrations between classes that was retained in the end²⁴.

<i>ratio max /min</i>	Premium		Reserves	
	CP2009	QIS 5	CP2009	QIS 5
1 Motor, third party liability	1.3	2.8	1.5	4.2
2 Motor, other classes	1.7	1.6	2.5	2.2
3 Marine, aviation and transport	1.7	1.4		3.0
4 Fire and other damage to property	1.3	1.8	1.8	2.6
5 Third-party liability		1.6	1.1	3.1
6 Credit and suretyship		1.2		1.0
7 Legal expenses		1.7	4.0	3.5
8 Assistance	1.3	1.3	7.3	3.9
9 Miscellaneous non-life insurance	7.8	2.2		1.5

Figure 11: dispersion between the results obtained by different methods and taken into account for the final choice

²⁴ Let us also highlight that this measure underestimates the “model’s margin of error”. Indeed, the only source of variation between the different methods adopted is, as indicated earlier, the number of degrees of freedom; but plenty of other parameters (such as the historical depth deemed pertinent or the choice of the underlying law of distribution enabling to transform volatility into regulatory VaR) are fixed here though they too could be subjects of diverging statistical assessments.

Given the current state of knowledge, Solvency II's aim to build a risk-based system appears utopian, since the margin of error associated with the risk estimations is much greater than the dispersion of risks between them. As such, it would seem more reasonable to build a fully *uncertainty-based* prudential system where capital requirements would be founded on (and at ease with) a qualitative, fixed assessment of the uncertainty of each LoB rather than such a supposedly risk-based system where capital requirements rely on a vain ambition of quantitative measurement.

Conclusion

One can discuss whether or not, in a context where issues concerning the protection of policyholders coexist with those concerning the development of insurance markets and the financing of the economy, the financial requirements imposed on firms should be based on prudential criteria alone. One could then discuss whether or not, from a strictly prudential standpoint, it is fully appropriate to base capital requirements on a risk measure. In any case, this was the double choice made by Solvency II. But it appears that the objective was not met: the calibrations on which capital requirements are based are not reliable. Let us stress that this study suggests Solvency II paradoxically led to a deterioration of resource allocation. Indeed, the allocation of capital resources has not been improved, and, in view of the energy expended on its conception and deployment, the allocation of human capital resources to which it has led is suboptimal.

Beyond this, one could question whether such a risk-based ambition can even be realized. We can't consider that risk measurement technologies don't currently exist, but the evidence presented in this article demonstrates that the issue is not a question of *refining* the precision – rather, it is a question of the current *inability* to characterize the relative risk of one activity in comparison to another. As such, we cannot rely on this basis for decision-making.

In this context, what should be done regarding the technical limitations and political trade-off issues?

Better identification of technical limitations

The study we have conducted can be enriched in three ways. First, by broadening the scope of the analysis from only premium and reserving risks to all risks.

Second, by not only testing internal credibility but also the calibrations' intrinsic margin of error through several sets of alternative inputs. While some studies of sensitivity to hypotheses have been conducted²⁵, it would now be useful to review these analyses in order to determine the associated margin of error, for each capital requirement component. We insist that the issue at stake is not so much sensitivities – which are forgotten as soon as they are presented and often serve as a discharge to present and use an unreliable result – as it is the *margin of error*; the sensitivity to a parameter should only be a means of determining it.

Third, by going one step further in the estimation of the overall margin of error of the SCR. Indeed, this paper focuses on the margin of error of some components of the SCR, considered separately. Articulating these measures with the study of the margin of error generated by the aggregation process of the different components, as explored in other papers (e.g. Puccetti and Rüschenendorf, 2012) would allow to determine the global margin of error of the capital requirements.

²⁵ And while they may be scattered, they are numerous (both internally and at aggregated levels for professional federations and regulators): the impact of the choice of such or such a yield curve reference, of the choice of such or such a calibration of market stress, of the choice of such or such a cat model rather than another, of such a management rule instead of another, of such a reference to ratings or to market spreads so as to evaluate a portfolio's sovereign default risk, etc.

Conscientious steering

Precisely, the sensitivity studies that were conducted were, to our knowledge, never meant to determine whether the measures were sufficiently precise to serve as a basis for decision-making. Still, they were intended to enable one hypothesis to be chosen over another. Three reasons can be put forward to explain this: first, the fact that a risk-based system was going to be implemented, so the issue was its characteristics rather than its merits; second, the modelers may have wanted to avoid devaluing their own work and thus risk being seen as useless²⁶; third, the diffuse and omnipresent feeling that it was “better than nothing” – that the result of a calculation is superior *in principle* to a fixed ratio.

We have shown that such a sentiment is unfounded at this stage. In the current state of knowledge, an insurance company’s capital requirement under Solvency II is the aggregation of several dozen components, for which it is not possible to know whether the relative orders of magnitude are being respected. The result is thus completely devoid of meaning; the implications of this lesson for current steering should be noted on both the macro-economic and the micro-economic levels.

At the macro-economic level, when these calibrations appear to be technically unreliable, they are no longer legitimate for establishing insurers’ capital requirements. This calls for an effort to rethink these requirements at a political level to articulate the whole set of impacted dimensions (e.g., financing of the economy, development of a market, etc.) beyond an artificial risk-based calibration. Of course, we could consider that the political issues have been

²⁶ Questioned by a consultant for a survey on ALM practices and strategic asset allocation methods, I had the following exchange:

- Me: The important thing is to know and to relate the margins of error.
- Him: Yes, you have to present the sensitivity analyses.
- Me: No, the sensitivities are anecdotal. What’s important is the margin of error.
- Him: But if we present the margins of error, the models will lose their credibility!

taken into account, as the CEIOPS seems to have censured itself by downward biasing its recommendations with respect to its technical estimations. But the CEIOPS, whose expertise is, theoretically, exclusively technical, does not have the legitimacy to make decisions and, indeed, does not seem to take responsibility for it. Nothing, to our knowledge, suggests that these choices have been steered conscientiously on the basis of a structured analysis.

At the micro-economic level, the main danger is for these tools to effectively be used as risk indicators in steering companies. The most precarious firms, thus those for which the quality of risk management is most sensitive, are all the more exposed because the pressure of a low solvency ratio will lead them to align their management with this framework in order to optimize the image they give to the regulator. Both in terms of discourse and implementation, with regard to use test requirements and risk management issues, it is important to dissociate the apprehension of risks from Solvency II's quantifications as much as possible. This would suggest leaving the SCR calculation and the solvency ratio management to the CFOs rather than to the CROs. Most of all, it calls for an increased place of the Pillar two (qualitative requirements) of Solvency II in the risk management framework of the companies, both in terms of dedicated internal resources and in terms of inputs for governance bodies. In particular, it highlights the importance of the ORSA in which the companies can address, for example, the issue of non-modelled risks such as the sovereign debt default. Alongside, this naturally calls for rebalancing the focus of the regulators, away from Pillar one, towards Pillar two.

References

ACPR (Autorité de Contrôle Prudentiel et de Résolution), (2011) “Solvabilité 2 : principaux enseignements de la cinquième étude quantitative d’impact (QIS5) ”, *Analyses et synthèses* n°1, 24p.

Alesina, A. and Tabellini, G., (2007) “Bureaucrats or Politicians? Part I: A Single Policy Task”, *American Economic Review*, **97(1)**, 169-179.

CEIOPS (Committee of European Insurance and Occupational Pensions Supervisors), (2006) “Quantitative Impact Study 2 Technical Specification”, *CEIOPS-PI-08/06*, 66p.

CEIOPS, (2007a) “QIS3 Technical specifications, Part I: Instructions”, *CEIOPS-FS-11/07*, 119p.

CEIOPS, (2007b) “QIS3 Calibration of the underwriting risk, market risk and MCR”, *CEIOPS-FS-14/07*, 53p.

CEIOPS, (2009) “Draft CEIOPS’ advice for Level 2 Implementing Measures on Solvency II: SCR Standard Formula Calibration of Non-Life Underwriting Risk”, *CEIOPS-CP-71-09*, 118p.

CEIOPS, (2010) “Solvency II calibration paper”, *CEIOPS-SEC-40-10*, 384p.

Cummins, J., & Phillips, R., (2009) “Capital adequacy and insurance risk-based capital systems”, *Journal of Insurance Regulation*, **28(1)**, 25-72, Fall 2009.

Danielsson J., (2002) “The emperor has no clothes: Limits to risk modelling”, *Journal of Banking and Finance*, **26(7)**, 1273-1296.

Danielsson J., (2008) “Blame the models”, *Journal of Financial stability*, **4(4)**, 321-328.

El Karoui, N., Loisel, S., Prigent, J.-L., & Vedani, J., (2015) “Market inconsistencies of the market-consistent European life insurance economic valuations: pitfalls and practical solutions”, *working paper*, online publication, <https://hal.archives-ouvertes.fr/hal-01242023>

Eling, M., & Holz Müller, I., (2008) “An overview and comparison of risk-based capital standards”, *Journal of Insurance Regulation*, **26(4)**, 31-60.

Embrechts, P., Puccetti, G. and Rüschendorf, L., (2013), “Model uncertainty and VaR aggregation”, *Journal of Banking & Finance*, **37**, 2750-2764.

EC (European commission), (2008) “Technical Specifications QIS4”, *MARKT/2505/08*, 293p.

EC (European Commission), (2010) “QIS5 Technical Specifications”, 330p.

EPC (European Parliament and Council), (2009) “Directive 2009/138/CE du Parlement européen et du Conseil du 25 novembre 2009 sur l'accès aux activités de l'assurance et de la réassurance et leur exercice (solvabilité II)”, *OJ L 335*, 1–155, 17.12.2009.

Harrison, M. & Samuel, S., (2011) “Two hundreds millions inputs. Can you trust risk weightings at European banks? ”, *Barclays Capital Equity Research*, 06.04.2011.

Kamega, A. & Planchet F., (2013) “Construction de tables de mortalité prospectives sur un groupe restreint : mesure du risque d'estimation, Bulletin français d'Actuariat”, *Bulletin Français d'Actuariat*, **13(25)**, 5-33.

Leroy, G. & Planchet, F., (2010) “Que signifie la ruine dans Solvabilité 2 ?”, *La Tribune de l'assurance*, 147, 55-56.

Pfeifer, D. and Strassburger, D., (2008) “Solvency II: Stability problems with the SCR aggregation formula”, *Scandinavian Actuarial Journal*, **1**, 61 – 77.

Pfeifer, D., (2013), “Correlation, tail dependence and diversification”, in C. Becker, R. Fried, S. Kuhnt (Hrsg.): *Robustness and Complex Data Structures. Festschrift in Honour of Ursula Gather*, 301 - 314, Springer, Berlin.

Puccetti, G. and Rüschendorf, L., (2012), “Bounds for joint portfolios of dependent risks », *Statistics & Risk Modeling*, **29**, 107–132.

Repulo & Martinez-Miera, (2014) “Comparing Flat and Risk-based Capital Requirements”, in *3rd European Banking Authority Policy Research Workshop*, London.

Sims, C., (2003) “Implications of rational inattention”, *Journal of Monetary Economics*, **50(3)**, 665–690.

Weber, R., & Darbellay, A., (2008) “The regulatory use of credit ratings in bank capital requirement regulations”, *Journal of Banking Regulation*, **10(1)**, 1-16.

Appendix

Successive risk calibrations proposed by the public authorities

The last two rows show, for each QIS, the value of indicators of dispersion between classes – they synthesize an order of magnitude for the measure’s dispersion capacity, which is to say the amplitude of the information supplied. The last two columns show, for each class, the value of indicators of dispersion between QIS – they synthesize an estimation of the amplitude of the noise interfering with the measure. The second indicator, for example, shows that for two-thirds of the LoBs, the calibration changes by a factor of two or more depending on the measure.

	2006	2007	2008	2010	2014	noise indicators	
	QIS2	QIS3	QIS4	QIS5	Preparatory phase	stdev	max-min/min
1 Motor, third party liability	5.0%	7.5%	9.0%	10.0%	10.0%	2.1%	1.0
2 Motor, other classes	12.5%	3.0%	9.0%	7.0%	8.0%	3.4%	3.2
3 Marine, aviation and transport	7.5%	5.0%	12.5%	17.0%	15.0%	5.0%	2.4
4 Fire and other damage to property	15.0%	10.0%	10.0%	10.0%	8.0%	2.6%	0.9
5 Third-party liability	10.0%	10.0%	12.5%	15.0%	14.0%	2.3%	0.5
6 Credit and suretyship	25.0%	12.5%	15.0%	21.5%	12.0%	5.8%	1.1
7 Legal expenses	10.0%	10.0%	5.0%	6.5%	7.0%	2.2%	1.0
8 Assistance	15.0%	10.0%	7.5%	5.0%	9.0%	3.7%	2.0
9 Miscellaneous non-life insurance	10.0%	12.5%	11.0%	13.0%	13.0%	1.3%	0.3
raw signal	Stdev	5.8%	3.2%	3.0%	5.4%	2.9%	
indicators	max-min/min	4.0	3.2	2.0	3.3	1.1	

Figure A3: calibrations associated with premium risk and associated raw signal and noise indicators

	2006	2007	2008	2010	2014	noise indicators	
	QIS2	QIS3	QIS4	QIS5	Preparatory phase	stdev	max- min/min
1 Motor, third party liability	15.0%	15.0%	12.0%	9.5%	9.0%	2.9%	0.7
2 Motor, other classes	15.0%	7.5%	7.0%	10.0%	8.0%	3.3%	1.1
3 Marine, aviation and transport	7.5%	15.0%	10.0%	14.0%	11.0%	3.0%	1.0
4 Fire and other damage to property	15.0%	12.5%	10.0%	11.0%	10.0%	2.1%	0.5
5 Third-party liability	10.0%	7.5%	15.0%	11.0%	11.0%	2.7%	1.0
6 Credit and suretyship	20.0%	15.0%	15.0%	19.0%	19.0%	2.4%	0.3
7 Legal expenses	20.0%	10.0%	10.0%	9.0%	12.0%	4.5%	1.2
8 Assistance	10.0%	15.0%	10.0%	11.0%	20.0%	4.3%	1.0
9 Miscellaneous non-life insurance	20.0%	10.0%	10.0%	15.0%	20.0%	5.0%	1.0
raw signal	Stdev	4.8%	3.3%	2.6%	3.2%	4.9%	

indicators	max-min/min	1.7	1.0	1.1	1.1	1.5
------------	-------------	-----	-----	-----	-----	-----

Figure A4: calibrations associated with reserving risk and associated raw signal and noise indicators

Chapitre 2 : La question de la pertinence des indicateurs

With-Profit Contracts Regulation and Procyclicality: A Comparative Analysis of Alternative Approaches

Introduction

In 2005, it was decided that Solvency II would come into effect in 2010 (ACAM, 2005; François, 2015); however, this was not the case until 2016, risking abandonment along the way in 2013. It was initially planned that there would only be a single QIS (Quantitative Impact Study) on the calibration of capital requirements; there were to be five in the end²⁷. These delays mainly resulted from the display of a particularly violent phenomenon of capital requirement procyclicality on the occasion of the 2007 financial crisis and the ensuing decline in interest rates – a phenomenon whose magnitude apparently had not been anticipated. This being a surprise is itself surprising. Indeed, as we will demonstrate, this procyclicality is the direct and inevitable consequence of the application to the most widespread insurance contracts of two combined principles intentionally placed at the heart of Solvency II: balance sheet economic valuation and risk-based capital requirements. Patches (countercyclical premiums, volatility adjusters, etc.) were developed between 2008 and 2015 to mitigate this procyclicality – never calling these two principles into question. These reduced the system’s legibility and, naturally, procyclicality remained, as its elimination is incompatible with the two core principles of Solvency II. This incompatibility and its consequences are the subject of the present paper.

In the banking sector, two kinds of mechanisms were identified to explain the procyclicality of the Basel regulations. First, available capital is reduced in unfavorable circumstances. When the assets fall, their economic value declines, and this accounting phenomenon reduces capital²⁸. Second, capital requirements increase in unfavorable circumstances. This phenomenon is a consequence of risk-based approach: as the quality of credit diminishes in a

²⁷ QIS 2 through 5 and LTGA, whereas only QIS 2 had been planned in 2005 (QIS 1 did not concern capital requirements).

²⁸ and increases leverage, thus leading to a reduction of activity and tending to increase the movement of asset prices on the markets in cases of liquidity stress and massive securities sales (IMF, 2008; Plantin et al., 2008, or, for a formulation using liquidity issues, Brunnermeier and Petersen, 2009, and, from an empirical viewpoint, Darolles et al., 2013).

crisis, this increases the risk measures and the related capital requirements (cf., Kashyap and Stein, 2004; Taylor and Goodhart, 2006).

To our knowledge, the principle of a prudential design that reflects a firm's net financial position and reduces capital in the event of an unfavorable situation is not debated. Only the magnitude and the speed of impact (increased in the case of valuation at market value) are debated (Plantin et al., op. cit.). The fact that capital requirements are procyclical is more debated, the main argument in favor of risk-based capital requirements being the incentive provided to decision-makers (Rochet, 2008), even if it means procyclicality is an implied consequence.

These now well-known mechanisms (see for example Auray and Gouriéroux, 2014, for a summary) are also at work in Solvency II, which introduces risk-based capital requirements and a prudential balance sheet at market value; however, they are greatly amplified by the combination of two other phenomena that are not found in the banking sector:

- the first is specific to insurance products: many retirement and savings guarantees are based on the profit sharing between insurers and policyholders, where the insurer guarantees a minimum return rate to the policyholder and shares with him the contingent surplus²⁹
- the second is specific to Solvency II: the risk-based approach has been implemented in a holistic way which takes into account the loss absorption capacity of the policyholders

What are the consequences of applying a Solvency II-type analytical framework to contracts built around a profit sharing mechanism?

When valuating assets and liabilities at market value ("economic" approach), expected future profits are integrated into the economic capital to assess the

²⁹ As we will underline later, one can notice, however, that our analysis of profit sharing between an insurer and its policyholders can be straightforwardly widened to another general profit-sharing mechanism : taxes, where a company shares the profits with the State and keep the losses.

company's net position (Frezal, 2016). At the same time, the risk-based approach retained in Solvency II consists in considering the impact of an adverse event with 0.5% probability *on economic capital* as the capital requirement (Laas and Siegel, 2016). This mechanism thus includes variations in expected profits into capital requirements.

When the initial situation is favorable – for example, when the company has accumulated unrealized gains or because the yields on the financial markets are high in comparison to the guaranteed rates of the liabilities – the insurer's expected profits are significant. In the event of a variation in the situation around this point, the insurer will share this profits variation with the policyholders. A given exogenous deterioration of the situation will thus reduce its expected profits to a limited degree, as the major part of the variation will be absorbed by policyholders. The impact of this deterioration on the insurer's economic capital will therefore be moderate and the capital requirements relatively low.

Let us now imagine the inverse case, where the initial situation is unfavorable – for example, because the company has unrealized losses or because the yields on the financial markets have fallen below the guaranteed rates of the liabilities. In this case, the insurer will have to honor the guaranteed rates whatever it takes. A given exogenous deterioration of the situation will reduce expected profits significantly, as the insurer will no longer be in a position of profit sharing with policyholders. The impact of this deterioration on the insurer's economic capital will be high and the capital requirements will be relatively high as well.

This double example describes the situation of European life insurance in the early 2000s, when Solvency 2 was conceived, on the one hand, and in the 2010s, as it enters into force, on the other. A product being launched during the former period generated economic capital (significant expected profits) and required little capital (potential degradation absorbed by policyholders). Today,

the same product no longer generates economic capital and it requires a significant amount of capital. This illustrates the fact that, from an economic view, a risk-based regulation is structurally procyclical when applied to profit sharing contracts (even in the case where the stress on assets does not increase in a crisis period), and that, in favorable periods, it does not encourage the creation of a capital buffer in anticipation of a slow deterioration of the financial environment.

Section 1 presents the contractual profit sharing mechanism specific to life insurance and the stylised model that allows us to describe the intuition of the mechanism at work under Solvency II-type regulation. Section 2 extends the results from the model to steering implications, including a comparison with the information and incentives provided by a non-economic and non-risk-based prudential mechanism. Section 3 provides a quantified experimental example.

1. With-profit contracts, intuition and modeling

1.1 Contract description

The prudential regulations most subject to institutional debates and academic analyses have to do with corporate law – specifically, the valuation of balance sheets and associated capital requirements. These are crucial aspects of prudential regulation, but they are only the tip of the iceberg. In fact, the most fundamental element in limiting an insurer’s bankruptcy risk is contract law: not making promises that are difficult to keep constitutes the basis of an insurer’s future capacity to fulfill its commitments³⁰.

There are two main types of contracts in life insurance: unit-linked policies and profit-sharing policies. Unit-linked are policies where the insurer does not guarantee capital, which obviously facilitates his ability to keep his promises. Profit-sharing policies are described, for example, in Ramlau-Hansen (1991) or

³⁰ This explains a number of exclusions in non-life contracts, such as damages not being covered in the event of war and the existence of guarantee cap (generally €100M in motor liability).

Norberg (1999); in a nutshell, these are policies where the insurer guarantees a return, nil or positive,³¹ capped by public authorities for prudential purposes. It is limited in France, for example, to a certain proportion of the average return on government bonds and, in practice, insurers have guaranteed a zero rate on new flows for several years. In Germany, a maximum rate for new flows was determined each year by the Ministry of Finance (1.25% in 2015) and, in practice, the rate proposed by the market corresponded to this theoretical cap. The direct corollary of such a cap is the biasing of the division of the wealth created in favor of the insurers. In order to limit this bias while still capping the rates guaranteed, regulations oblige insurers to pay out a certain share of their profits to policyholders. In France, for example, life insurers must pay policyholders at least 85% of financial results (and 90% of underwriting profits). Likewise, when Germany lowered the guaranteed rate ceiling from 1.75% to 1.25% in 2014, it simultaneously increased the minimum rate of underwriting results to be paid out from 75% to 90%.

Even when it is not imposed by regulations as is the case in France and Germany, this prudential mechanism which combines a low guaranteed rate with profit sharing may be used voluntarily by companies for commercial reasons (*with-profit policy* in the UK, *participation policy* in the US). These profit sharing policies represent over 80% of the €1,600B in life insurance reserves in France (ACPR, 2015) and more than 90% of retirement and savings reserves in Germany (EIOPA, 2014)³².

1.2 Modeling of the contract and associated profit function

The methods of balance sheet valuation (Bonnin, Juliard and Planchet, 2014) and of risk measurement (Thérond & Planchet, 2007) are then merely grafted

³¹ In Europe, for example, insurance reserves with capital guarantees carry a guaranteed rate of 3% on average (EIOPA, 2014).

³² Non-unit-linked contracts (i.e., adapted to profit sharing) represented 80% of retirement and savings reserves in Europe at the end of 2013 (EIOPA, 2014). On average, at the end of 2013, they represented over 70% of life insurance reserves in OECD countries for which information was available (OECD, 2015). This second percentage should be treated with caution, though, as it is calculated from figures that appear to be very fragmentary even in the countries where they are available.

onto this contractual reality in order to provide an aggregate financial assessment at the company level.

To model the intuition, we consider a one-period model and stylize the profit sharing mechanism through its financial component. Let us consider a contract with two characteristics:

- a minimum guaranteed rate (MGR) that must be credited to the policyholder;
- a profit sharing rate, $1 - \alpha$, that corresponds to the share of profits allocated to the policyholder when profits are realized, with $\alpha \in [0; 1]$ corresponding to the share of profits retained by the insurer.

At the beginning of the period, the insurer invests the premium and will obtain at the end of the period a rate of return r_f . Its corresponding profit π_f can be written as follows:

$$\pi_f = \begin{cases} \alpha (r_f - MGR) & \text{if } r_f \geq MGR \\ (r_f - MGR) & \text{if } r_f < MGR \end{cases} \quad [1]$$

from which we may obtain *Figure 1*, illustrating the insurer's profit depending on the return obtained over the considered period:

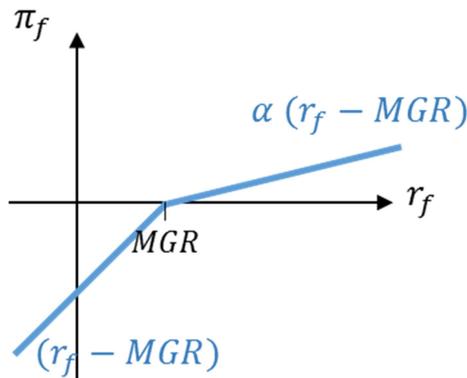


Figure 1: insurer's profit depending on the asset return

1.3 Valuation of economic capital

The effective return r_f observed at the end of the period can be written as:

$$r_f = \tilde{r}_f + \varepsilon, \text{ with } E(\varepsilon) = 0,$$

where \tilde{r}_f is the expected return³³ at the beginning of the period, depending on both exogenous and industry-wide parameters, such as the level of the rate curve, and endogenous and peculiar parameters, such as the aggressive or prudent past policy of realizing gains or losses and the resulting amount of accumulated unrealized gains.

At the beginning of the period, the insurer has an amount of basic own funds K_0 . The economic own funds (EOF) take into account the basic own funds and the expected profits and may thus be written as:

$$EOF = K_0 + E(\pi_f | \tilde{r}_f) \quad [2]$$

The expected return \tilde{r}_f is thus an indicator of the economic and financial situation, more or less favorable to the insurer (either due to exogenous and industry-wide parameters, or due to endogenous and peculiar reasons, as mentioned): the higher \tilde{r}_f is, the greater the expected profits are, the healthier the insurer's financial situation and its economic own funds will be.

Property 1:

$E(\pi_f | \tilde{r}_f)$ is an increasing concave function of \tilde{r}_f which asymptotes in \tilde{r}_f are

$$\begin{cases} \text{at } +\infty, & \alpha(\tilde{r}_f - MGR) \\ \text{at } -\infty, & (\tilde{r}_f - MGR) \end{cases}.$$

Proof:

It is an increasing function of \tilde{r}_f as it is the combination of a continuous linear function (Expectancy) and an increasing function ($\pi_f(\tilde{r}_f)$). It is a concave function as it is the combination of this continuous linear function and this concave function. Its asymptotes are

³³ In a financial context where the discounted price process is a martingale, we would have $\tilde{r}_f = r_0$, where r_0 is the effective market rate of return when selling the contract. In an insurance context where it is possible to have stored unrealized profits and losses or to obtain an underwriting result, \tilde{r}_f may differ from r_0 .

obtained directly, once ε is integrable, by writing $E(\pi_f | \tilde{r}_f)$ in the form: $E(\pi_f | \tilde{r}_f) = \alpha(\tilde{r}_f - MGR) + (1 - \alpha) E(\min(\tilde{r}_f + \varepsilon - MGR; 0) | \tilde{r}_f)$.

Figure 2 thus illustrates the expected value of future profits according to the expected central rate of return \tilde{r}_f (Figure 3):

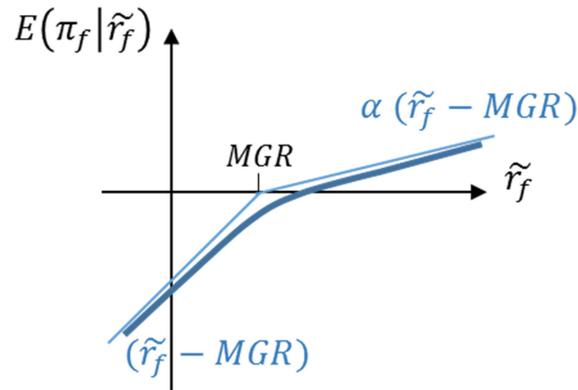


Figure 2: expected value of future profits depending on the initial situation

1.4 Risk-based capital requirement and procyclicality

When selling the contract, the insurer calculates its Solvency Capital Requirement (*SCR*). This risk-based requirement is designed to allow the insurer to resist a stress of a given probability: it corresponds to the fall in economic own funds in the eventuality of this stress. It can therefore be written as:

$$SCR = EOF - EOF'$$

where *EOF'* corresponds to the recalculation of capital in the virtual stressed situation considered.

Otherwise put, s being the magnitude of the applied stress (annual VaR at 99.5% in the context of Solvency II³⁴) and $\tilde{r}_f' = \tilde{r}_f - s$ the expected central rate of return in a stressed situation, we get:³⁵

$$SCR = E(\pi_f | \tilde{r}_f) - E(\pi_f | \tilde{r}_f') \quad [3]$$

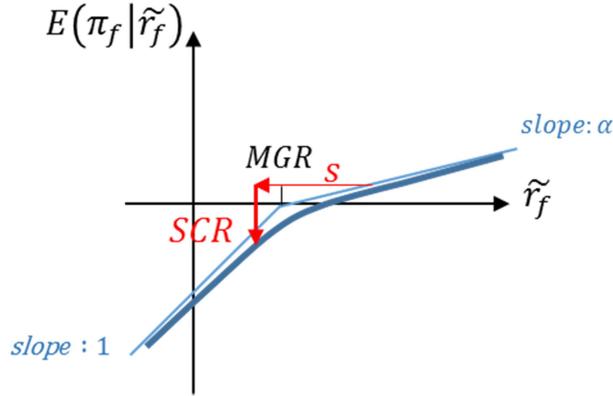


Figure 3: capital requirement

Property 2:

Thus, if a capital requirement is (i) risk-based relying on a constant stress³⁶, (ii) based on economic capital, and (iii) applied to profit sharing contracts, then the more favorable the insurer's initial financial situation, the lower the Solvency Capital Requirement is: if \tilde{r}_f is high, the capital requirement is low, and inversely, if \tilde{r}_f is low, then the capital requirement is high.

³⁴ Under Solvency II, a set of stresses (movement of the interest rate curves, fall of equity and property markets, etc.) are to be considered, and their impact to be measured. Afterwards, the impact of an aggregate stress, of lower magnitude in each component, is to be assessed.

³⁵ In order to simplify the formalization, we suppose that K_0 is not affected by a stress. This is actually the case if the assets are ring-fenced and those backing basic own funds are invested in cash. For firms that do not manage their assets in this way, taking into account the impact of the stress on K_0 translates the results obtained without modifying its structure.

³⁶ Regarding credit risks, following the habits of banking regulation, the deeper the crisis, the higher the stress, which amplifies the highlighted mechanism. Regarding interest rate risk or equity risk, EIOPA tried to mitigate the procyclicality by reducing the stresses in unfavorable times. As we will demonstrate in the empirical example, the amplitude of the reduction in stress is far from fully preventing the procyclical effect.

More precisely, the capital requirement evolves between a floor and a cap, whose ratio corresponds to the profit rate α for profits retained by the insurer, as illustrated in Figure 4:

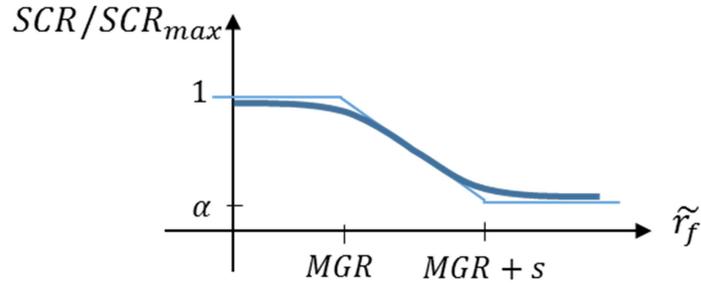


Figure 4: impact of the initial situation on the capital requirement

Proof:

Due to the concavity of $E(\pi_f | \tilde{r}_f)$ (Property 1), it results straightforwardly from equation [3], that, at fixed s , SCR is a decreasing function of \tilde{r}_f , as illustrated in Figure 5:

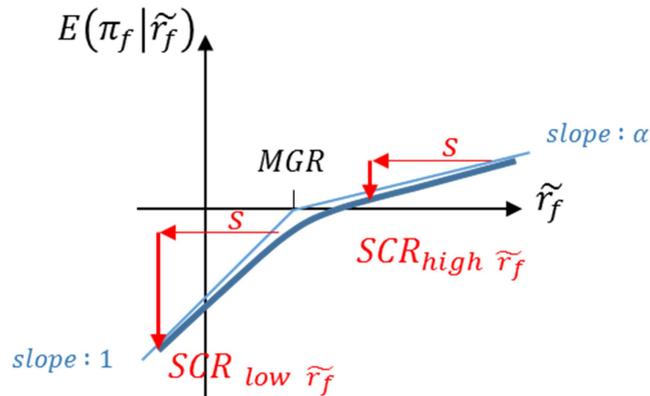


Figure 5: impact of the initial situation on capital requirement

Ratio α between the cap and the floor of the SCR results straightforwardly from the asymptotic profile of $E(\pi_f | \tilde{r}_f)$ (Property 1).

Thus, the combination of the three characteristics *i*, *ii*, and *iii* from Property 2 inevitably leads to a procyclical regulation even if the magnitude of the

exogenous stress does not increase in an unfavorable situation and simply remains constant.

2. Comparative steering insights of risk-based, economic (RBE) system vs. non-RBE system

In a prudential system that relies on economic valuation, expected profits generated by a product cover a part of its capital requirement. As a result, the minimum raised capital to be obtained from (or that cannot be released to) the shareholders differs from the capital requirement: we now turn our attention to the minimum raised capital and its derivative.

We will now focus on:

- (1) The minimum raised capital itself – that is, the gross capital requirement (the SCR in Solvency II) *net of economic capital generated by the product* (i.e., net of future profits)– an amount that constitutes the actual denominator in the profitability ratio of the shareholders³⁷ investments and thus also constitutes the information on which a rational actor bases his investment and risk/return arbitrage decisions (whereas the capital requirement does not).
- (2) The sensitivity of this minimum raised capital to an adverse event – a sensitivity that directly affects the company’s ability to distribute dividends (or its need to raise capital) in the short term, and thus affects management decisions.

This will allow us to assess the operational incentives provided to managers. We will compare the evolution of these two indicators according to the firm’s financial position,

- in a risk-based economic environment (RBE) such as Solvency II – for which we have described the impact on the capital requirement – on the one hand,

³⁷ E.g. to the holding, in the case of a subsidiary.

- and in a prudential environment such as Solvency I, which is neither risk-based nor economic (non-RBE) – where the capital requirements are constant and only expected losses can affect the available capital – on the other hand.

2.1 Minimum raised capital and anticipation of slow risks

Non-RBE environment

In a Solvency I-type case, (i) capital requirements are flat and (ii) the balance sheet is not measured from an “economic” view but rather from a prudent accounting view: when expected profits are positive, they are *not* taken into account, and when they are negative, they *are* taken into account. Indeed, when the return \tilde{r}_f falls below the minimum guaranteed rate MGR , the difference must be reserved: it decreases the available own funds³⁸ and increases the raised capital that the shareholder must immobilize.

This leads to the non-RBE curve in Figure 6.

RBE environment

Returning to the notations above, the minimum raised capital K (which K_0 must not fall below) that the shareholder must immobilize in order for the company to be able to operate is:

$$K = SCR - E(\pi_f | \tilde{r}_f) = -E(\pi_f | \tilde{r}_f') \quad [4]$$

This minimum raised capital can be represented according to the initial economic and financial situation \tilde{r}_f , such as we see in Figure 6:

³⁸ Formally, expected losses are not taken into account as negative capital, but the prudence requirement leads to their being reserved (*via* the mechanisms of *provision pour aléa financier*, *provision pour risque de taux* or *provision globale de gestion* in French accounting, for example – cf. Frezal 2015). The negative result corresponding to the setting up of these reserves generates a reduction of share capital that needs to be compensated for by the shareholder. This then corresponds *de facto* to an attribution of expected losses to the capital.

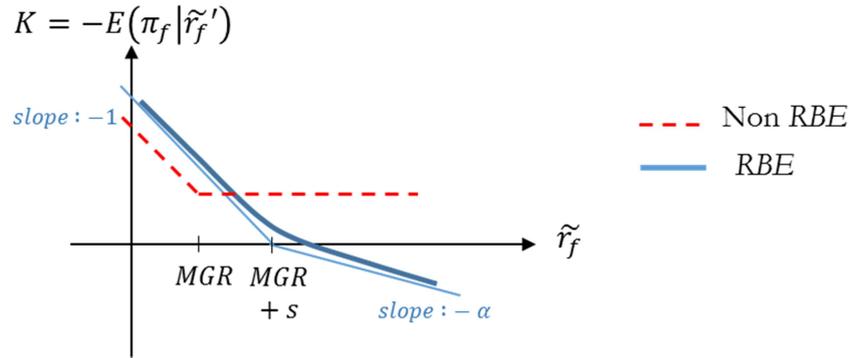


Figure 6: minimum raised capital depending on the initial situation

Property 3:

In an RBE environment:

1. *If the economic and financial situation is very favorable, the shareholder can completely avoid immobilizing his capital.*
2. *Inversely, when the initial financial situation degrades, the amount of capital that must be immobilized by the shareholder increases, and there is no limit to this amount.*

Proof:

Point 1 follows from Equation 4 in the way that, if in the worst case envisioned by the regulations, the product remains profitable for the insurer – that is, asymptotically³⁹, when the expected post-stress return is higher than the minimum guaranteed rate ($\tilde{r}_0 - s > \text{MGR}$), then the raised capital is negative. Indeed, expected profits are such that the economic capital created by the product is greater than the capital requirement.

When the situation is degraded, the limitless character of the raised capital to be immobilized (point 2) results from the combination of two elements:

1. *beyond a certain level, the capital requirement (SCR) is asymptotically limited (property 2);*

³⁹ To place oneself in the asymptotic situation is to neglect the option costs (TVOG – time values of option and guarantees) corresponding to the risk ε .

2. *but if the situation degrades, then the expected value of future profits continues to degrade and the shareholder must compensate for it with more capital.*

In a degraded situation, the profile of the minimum raised capital is thus analogous to that observed in a non-RBE environment, with only the level being translated.

Consequences

In a non-RBE system, the shareholder must immobilize capital even in favorable situations. For example, under Solvency I, shareholders had to immobilize capital for an amount corresponding to 4% of reserves.

In an RBE system, let us consider, for example, a policy with a minimum guaranteed rate of 3% (average MGR on the European market), associated with a return stress s of 1% (magnitude of the downward interest rate stress under Solvency II). In 2005, the return on bond portfolios was 5%, so a post-stress scenario would still have generated profits: such a policy would have created more capital than it consumed – the insurer would have been able to release all of its share capital and the shareholder to obtain an infinite return⁴⁰. Ten years later, when the rates have fallen, the stock of contracts accumulated by the insurer becomes extremely capital intensive, requiring strong injections from the shareholder and sending a radically different message regarding the product's level of risk.

This characteristic can thus be interpreted as short-sightedness on the part of the regulation: in favorable periods, the signal sent – focused on the 1-year horizon stress of a given magnitude s – is artificially optimistic and does not generate the right incentives to anticipate the risk of a long-term collapse of the economic and financial context.

⁴⁰ By “infinite”, we mean non capped, as the denominator of the ratio can be as low as wanted.

2.2 Sensitivity of the minimum raised capital and anticipation of short-term risks

Beyond the level of minimum raised capital that the shareholder needs to immobilize, the sensitivity of this amount to a potential loss is an important criterion for steering. Indeed, it is a signal that incites the insurer to adopt a strategy that is more or less exposed to volatility.

We normalize the portfolio size to 1 in order to assess the magnitude of an event by its impact $-dr$ on the return, and note its impact dK on the minimum raised capital.

Non-RBE environment

In the case of a non-RBE environment, such as Solvency I, there are two possible scenarios:

- If the expected profits are positive ($\tilde{r}_f > MGR$), then they are not taken into account and $dK(\tilde{r}_f) = -dSCR(\tilde{r}_f) = 0$,
- If the future profit prospects are negative ($\tilde{r}_f < MGR$), then they are taken into account and $dK(\tilde{r}_f) = -dSCR(\tilde{r}_f) + dE(\pi_f|\tilde{r}_f) = dE(\pi_f|\tilde{r}_f) = dr$, with the entirety of the stress being supported by the shareholder.

RBE environment

We neglect the option costs here to place ourselves in the asymptotic situation.

In an RBE environment such as Solvency II, it follows from [4] that:

$$dK(\tilde{r}_f) = E(\pi_f|\tilde{r}_f') - E(\pi_f|\tilde{r}_f' - dr)$$

Two possible cases may thus follow from Property 1:

- If the initial financial situation is favorable ($\tilde{r}_f' > MGR$, or $\tilde{r}_f > MGR + s$), then $dK(\tilde{r}_f) = \alpha \cdot dr$, with the stress being absorbed significantly by the policyholders,

- If the initial financial situation is not favorable ($\tilde{r}_f' < MGR$, or $\tilde{r}_f < MGR + s$), then $dK(\tilde{r}_f) = dr$. Thus the entirety of the stress is supported by the shareholder⁴¹.

These sensitivities are represented in Figure 7:

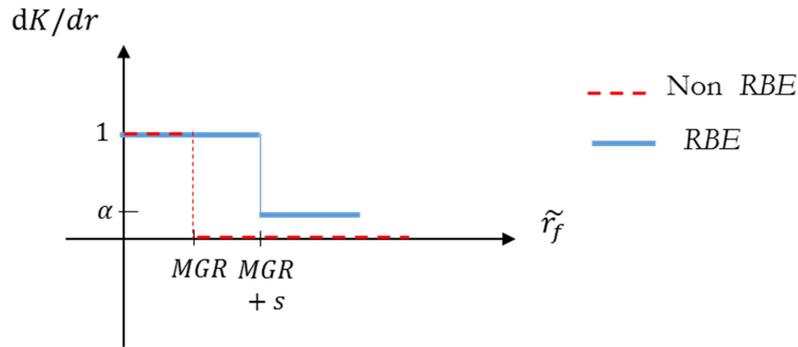


Figure 7: sensitivity of the minimum raised capital to a loss

Consequences

From this standpoint, an RBE-environment provides an improvement in terms of the ability to anticipate in comparison to a non-RBE one. Indeed, in an RBE system, the sensitivity to profit variations increases as soon as the expected return falls below an alert threshold $MGR + s$ – that is, before the firm begins to realize losses. As a result, this sensitivity acts as an early warning apt to lead the firm’s management to adopt a behavior that would reduce the firm’s exposure to volatility. Conversely, in the case of a non-RBE system, the sensitivity to economic fluctuations increases only when the expected return falls below MGR – that is when the firm begins to realize losses and is fully subject to any change in returns.

⁴¹ Either when the rate of return falls beneath the minimum guaranteed rate because the entirety of the loss is assigned to the capital that must be reconstituted (zone $\tilde{r}_f < MGR$ of figure 4a), or when the rate of return remains higher than the minimum guaranteed rate because the assignment of a portion of the loss to policyholders *via* the decrease of profit sharing is compensated in the insurer’s perspective by the increase of the capital requirement SCR (zone $\tilde{r}_f \in [MGR ; MGR + s]$ of figures 6 and 7).

3. Implementation within Solvency II

We now present a simulation that evaluates the capital requirement and the minimum raised capital for a classic retirement portfolio so as to illustrate the orders of magnitude associated with this procyclicality within the framework of Solvency II.

3.1 The characteristics of the simulation

Retirement is both, (i) from a technical viewpoint, particularly exposed to interest rate risk due to the difficulty of backing long-term liabilities with the assets available on the markets, and, (ii), from a political perspective, currently under consideration as regards the pertinence of applying Solvency II to it⁴².

We consider distinctly two retirement products with the following classic characteristics:

- 2% and 3% MGR;
- 90% profit sharing rate (Borel-Mathurin et al., 2018);
- and a 25-years duration.⁴³

These liabilities are backed by an asset allocation which characteristics are the following:

- 100% fixed-rate bonds;
- Maturities of 3, 8 and 11 years for corporate bonds and of 5, 10, 15, 20, 25 and 30 years for government bonds;

⁴² Pension funds have been excluded from the scope of the Directive (*European Parliament*, 2015) and certain legal structures such as IORPs are subject to transitional measures deferring their submission to Solvency II (ACPR, 2014, or *European Parliament and Council* 2015, §62).

⁴³ These parameters were chosen to be representative of a retirement portfolio launched in the 90's and therefore receiving premiums in the 2000's.

The details of the portfolio are the following:

- 68,315 working-age adults, 40 years old on average having an average deferred life annuity of €500,
- 15,099 pensioners, 67 years old on average having an average immediate life annuity of €1,389,
- 5,370 beneficiaries, 68 years old on average having an average immediate life annuity of €1563.

The formula used is the following: $Duration = \frac{\sum_{i=1}^{100} i \times Flow_i}{\sum_{i=1}^{100} Flow_i}$

- 10-years overall duration;⁴⁴
- Purchased at par at the beginning of the projection.

The asset-liability interactions as well as the associated profits and losses are modelled with a 50-year horizon.

Solvency II identifies a series of stresses, measures the impact of each on capital according to the methodology presented above, and then aggregates them through a correlation matrix. The calculations performed are based on the Standard Formula determined by EIOPA. However, we have only simulated the three main components:

- interest rate risk (the main component of market risk)⁴⁵,
- longevity risk (the main component of underwriting risk),
- and spread risk (the main component of credit risk).

For a similar retirement portfolio observed in a major firm, the interest rate risk alone represented 74% of the total aggregated SCR at the end of 2011.

Five interest rate environments have been simulated, corresponding to the yield curves provided by the EIOPA for 2015, 2012, 2009, 2006, and reconstituted for 2002 (cf. Figure 8). To present the results, we consider for each curve the average coupon rate of the portfolio of assets.

⁴⁴ The coupon frequency is 1 per year, and the detailed amounts are the following:

Bond type	Nominal (€)	Maturity (yrs)
corp. fix. rate	100 159 216	3
corp. fix. rate	200 318 431	8
corp. fix. rate	300 477 647	11
gov. fix. rate	154 091 101	5
gov. fix. rate	247 452 180	10
gov. fix. rate	257 150 855	15
gov. fix. rate	260 413 961	20
gov. fix. rate	164 968 120	25
gov. fix. rate	117 834 371	30

⁴⁵ For each year, the stresses retained are those provided by the EIOPA (estimated for 2002). This experimental case incorporates a mitigating effect, as the downside stresses are not constant, but are weaker the lower the yield curve is, which has a countercyclical impact. Further on, we will observe that despite this measure the procyclicality identified through the stylised model still appears when considering the aggregation of the different risk components.

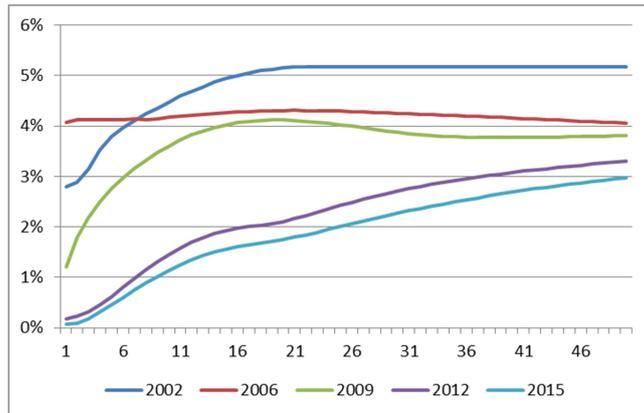


Figure 8: Yield curves used

3.2 Capital requirements

We first present the capital requirement of the portfolio. In order to have a point of comparison, we consider the ratio between capital requirement and mathematical reserves; under Solvency I, this value was constant and set at 4%. The results obtained are presented in Figure 9.

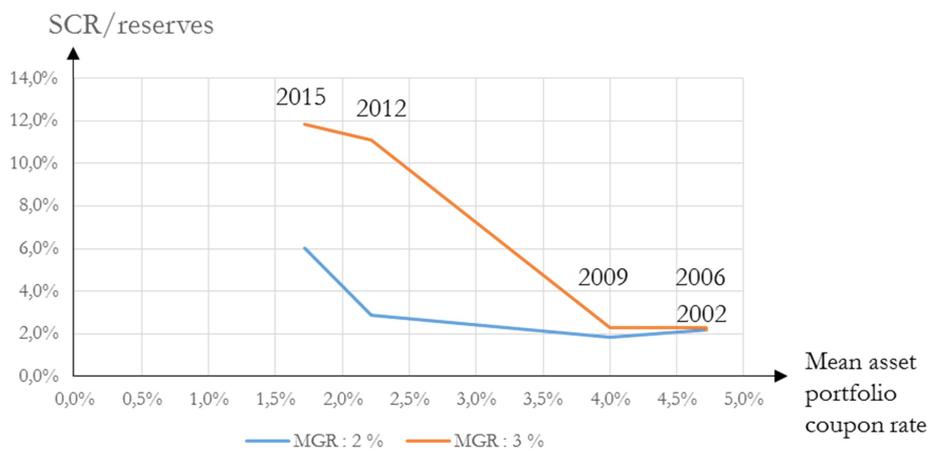


Figure 9: Capital requirement, retirement policy

We observe an inverted and flattened S-curve that goes from a cap (when the asset's rates of return are low) to a floor (when the rates are high), fully consistent with the figures anticipated by our stylised model. The entire curve

of the portfolio with a 3% MGR is described between 2002 and 2015. For the portfolio with a 2% MGR, the curve is translated toward low rates and only the bottom part of it is seen, as the coupon rates obtained in 2015 were still high enough to prevent the observation of SCR saturation⁴⁶.

3.3 Minimum raised capital

Here again, we consider the ratio between the minimum raised capital and the mathematical reserves. The results obtained are presented in Figure 10:

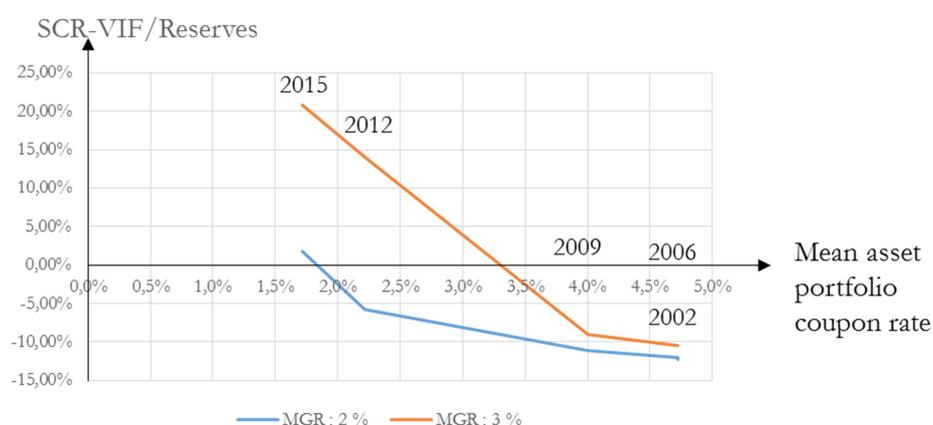


Figure 10: Minimum raised capital, retirement policy

Once again, the ratio corresponds to the development forecasted by the theoretical model presented earlier: when rates of return are low, the minimum raised capital is very sensitive to the economic situation, and when returns are high, the minimum raised capital does not depend much on the asset's return. Naturally, this curve is translated to the left (low rates of return) for lower MGR.

⁴⁶ The slight increase in the curve for the portfolio with a 2% MGR when the rates are higher can be explained by the fact that, with the specific combination of parameters retained, the portfolio would be exposed to an increase in the yield curve (decrease of the assets' market value) rather than to a decrease in the yield curve (decrease of the return on reinvestments). This is particularly contingent on the characteristics of the liability portfolio, the balance sheet's asset/liability backing, the prospective ALM management rules retained, and the calibration of stresses that may not be constant.

We may note that if Solvency II had been in effect in 2005, a product that is currently considered very risky would have required raising a negative amount of capital⁴⁷. This confirms that when the situation is favorable, such regulation sends a message that products are not risky and it is unnecessary to hold capital. There is reason to fear that if Solvency II had been in effect at this time, it would have led insurers to reduce their capital on a massive scale⁴⁸, as prospective future profits combined with low capital requirements would generate self-financing. During the financial crisis, then, they would have had to claim bankruptcy or raise huge amounts of capital, at the same time as the banks themselves needed to bolster their capital base. Similarly, the prospect of sharp increases in capital needs when Solvency II fully comes into force can only be cause for concern.

As such, an insurance company that steers its activity (e.g., its product launch decisions) on the basis of the minimum raised capital, provided by a Solvency II-type regulation, would be short-sighted. Such regulation does not provide companies with information on the slowly emerging risks that may materialize over the course of a product's lifecycle. Thus, it may be difficult for insurers to consider Pillar 1 of Solvency II as an internal risk management tool.

Conclusion

The combination of an “economic” valuation of an insurer's balance sheet and a risk-based regulation will thus inevitably be procyclical – particularly for profit sharing contracts. This significantly reinforces risk-based regulations' other sources of procyclicality, which are already well-known. Furthermore, the mechanism we have just described for contracts where an insurer shares profits with policyholders applies more broadly, and just as directly (though with a

⁴⁷ Let us remember, though, that we have limited ourselves to the calculation of the SCR components most important to the retirement activity. Considering all components of the SCR (expense risk, operational risk, etc.) would lead to an increase in the minimum raised capital and to moderate the sweep of this observation.

⁴⁸ In 2005, insurers had already adopted an aggressive capital management policy: basic own funds life insurers (in France) accounted for around 5% of reserves (ACAM, 2006), which was only slightly higher than the 4% regulatory requirement.

lower α), to another common profit sharing mechanism: that between a financial institution and the State in the context of corporation tax, where the financial organization bears all losses in an unfavorable situation and shares profits in a favorable one.

The underlying phenomenon generating the procyclicality of *capital requirements* seems poorly understood even now. The case of Solvency II's implementation tends to suggest that this phenomenon had not been anticipated by its designers (regulators and industry professionals), since the counter-cyclical measures only began to be studied once this procyclicality had materialized. The fact that the measures designed to fight against procyclicality seek more to patch up problems than to revise the regulation's fundamental characteristics may be due to the absence of theory underpinning this mechanism (and thus the absence of understanding of the inevitability of this procyclicality).

Furthermore, the analysis of the *effectively* minimum raised capital demonstrated that, when the context is favorable, a Solvency II-type regulation does not furnish the right information for managing the risk of gradual collapse and degrades the situation by creating an erroneously positive perception of certain contracts. Thus, if Solvency II had been in effect in 2005, retirement products now considered to be very risky would have been generative of capital, the insurance companies could have massively released their capital, and the state of the European insurance industry would doubtless be far more degraded now than is the case; moreover, during the crisis, the banks would have been confronted with huge capital raising operations in competition with their own.

Notwithstanding, we cannot conclude that an RBE regulation such as Solvency II is structurally less appropriate for managing risks than a non-RBE regulation such as Solvency I. Indeed, *sensitivity to losses of the raised capital* is also a valuable regulatory tool, as it provides an incentive for steering risks. In the framework of Solvency II, this sensitivity increases when the situation comes close to losses; in Solvency I, the sensitivity only increases when the company actually

begins to realize losses. From this point of view, Solvency II provides an early warning.

In terms of recommendations, four points come into view:

- The optimization of the design of incentives produced by prudential regulation cannot be reduced to a consideration of the capital requirement. The minimum raised capital and its sensitivity are equally crucial to steering, and thus so are incentives provided to managers.
- Risk assessments cannot be left to a Solvency II-type calculation of capital requirements: only stress tests considering a longer temporal horizon that corresponds to the lifecycle of products are able to capture the associated risks.
- In order to avoid the perverse effect of being able to release all of one's capital in good times, it would be advisable to set a floor for the minimum raised capital. This remains compatible with the reliance on a Solvency II-type capital requirement and benefits from the associated early warning effect.
- If we believe that a regulation should be risk-based, the risk measure would benefit from being based on the gross guarantee provided (e.g., the MGR) rather than on prospective future profits. Whatever would be lost in conceptual elegance would be gained in the avoidance of perverse effects (e.g., the procyclicality described here, as well as the heterogeneity of prospective future profits models and the resulting lack of a level playing field).

References

ACAM, (2006) : *“Rapport 2005”*, Autorité de Contrôle des Assurances et des Mutuelles, Paris, 103p.

ACPR, (2014) : *“Focus Solvabilité II, Numéro spécial Omnibus IP”*, Autorité de Contrôle Prudentiel et de Résolution, Paris, 21p.

- ACPR, (2015) : “*Les chiffres du marché français de la banque et de l’assurance 2014*”, Autorité de Contrôle Prudentiel et de Résolution, Paris, 140p.
- AURAY, S., GOURIEROUX C. (2014) : “The Procyclicality of financial markets regulation”, *Labex Louis Bachelier, Opinions et Débats*, n°4, April, 64p.
- BOREL-MATHURIN, F., DARPEIX, P.E., GUIBERT, Q., LOISEL, S., (2018) : Main Determinants of Profit Sharing Policy in the French Life Insurance Industry, *The Geneva Papers on Risk and Insurance - Issues and Practice*, Vol. 43(3), 420-455.
- BONNIN F., JUILLARD M., PLANCHET F. (2014) « Best estimate calculations of savings contracts by closed formulas - Application to the ORSA », *European Actuarial Journal*, Vol. 4, Issue 1, Page 181-196.
- BRUNNERMEIER, M., PETERSEN L. (2009) : “Market Liquidity and Funding Liquidity” », *Review of Financial Studies*, 22, pp. 2201-2238.
- DAROLLES, S., GAGLIARDINI, P., GOURIÉROUX, C. (2013) : “Survival of Hedge Funds : Frailty vs. Contagion”, *Univ de Lugano, DP.*, 71p.
- EIOPA (2014) : “*EIOPA Insurance stress tests 2014*”, European Insurance and Occupational Pensions Authority, Frankfurt, *EIOPA-BOS-14-203*, 244p.
- EPC (2014) : “*Directive 2014/ 51/ EU of the European Parliament and of the Council*”, European Parliament and Council, Brussels, OJ L 153, 22.5.2014.
- EP (2015) : “*Occupational pensions, Revision of the Institutions for Occupational Retirement Provision Directive (IORP II). Briefing EU Legislation in Progress*” European Parliament, Brussels, 11.12.2015, 12p.
- FRANÇOIS P. (2015) : “La fabrique sociale d’une *tabula rasa* : le lancement de *Solvency II*”, *Working paper PARI*, 61p.
- FREZAL, S. (2016) : “De quoi Solvabilité 2 est-il le nom ?”, *Working paper PARI*, 56p.

- IMF (2008) : “*Containing Systemic Risks and Restoring Financial Soundness, Global Financial Stability Report*”, International Monetary Fund, Washington, April, 211p.
- KASHYAP, A., STEIN, J. (2004) : “Cyclical implications of the Basel II capital standards”, *Economic Perspectives*, Q1, pp. 18-31.
- LAAS, D., SIEGEL, C. F. (2016) : “Basel III Versus Solvency II: An Analysis of Regulatory Consistency Under the New Capital Standards”, *Journal Risk and Insurance*, doi:10.1111/jori.12154, pp. 1231-1267.
- NORBERG, R. (1999) : “A theory of bonus in life insurance”, *Finance and Stochastics*, 3(4), pp. 373-390.
- OECD (2015) : “*OECD Insurance statistics, 2006-2013*”, Organisation for Economic Co-operation and Development, Paris, 207 p.
- PLANTIN, G., SAPRA, H., SHIN, H. (2008) : “Marking to market: Panacea or Pandora’s Box”, *Journal of Accounting Research*, 46, pp.435-460.
- RAMLAU-HANSEN, H. (1991) : “Distribution of surplus in life insurance”, *Astin Bulletin*, 21, pp. 57-71.
- ROCHET, J.C. (2008) : “Procyclicité des systèmes financiers : est-il nécessaire de modifier les règles comptables et la réglementation actuelles ?”, Banque de France, *Revue de la stabilité financière*, n° 12, October, pp. 105-110.
- TAYLOR, A. AND GOODHART, C. (2006) : “Procyclicality and volatility in the financial system: the implementation of Basel II and IAS 39”, In: Gerlach S. and Gruenwald P. (eds) *Procyclicality of Financial Systems in Asia*. Palgrave Macmillan, London, pp. 9-37.
- THEROND P.E., PLANCHET F. (2007) « Provisions techniques et capital de solvabilité d'une compagnie d'assurance : méthodologie d'utilisation de Value-at-Risk », *Assurances et gestion des risques*, Vol. 74 (4).

Partie II : Dysfonctionnements des mesures de risque dans la régulation des assurances : proposition d'un cadre théorique et applications au coût du risque

Chapitre 3 : Les défauts de l'approche statistique des risques : proposition de théorisation et application à Solvabilité 2

Risk and statistics: a prudential regulation fallacy

“Not only is it not right, it isn’t even wrong.”

W. Pauli

Introduction

For more than two centuries now on, insurers and their actuaries successfully use statistics to run their businesses. As a consequence, a belief emerged: statistics and actuarial tools are appropriate to manage risk and make decisions when facing uncertainty. This belief has become popular in recent decades, and implemented in various types of regulations. In the European Union (EU), the Solvency II Value-at-Risk (VaR)-based capital requirements have been considered cutting edge regulation to improve prudential supervision and enterprise risk management (ERM). I assert that this use of statistics goes beyond their actual field of appropriateness: regulations such as these are based on a muddling of the tools used for managing risk and those used for steering business profitability.

Insurers are frequently presented as the professionals of risk management: it is commonly considered that their core business is to “manage risk”. Of course, insurers are subject to inherent business uncertainty that must be managed: They are exposed to operational risks, to juridical changes that may increase the level of claims, to market risks such as a drop of equity prices, etc. In this matter, they are not different from any other industry: An oil company, for example, is subject to operational risks, to juridical changes regarding environmental laws, to market risks such as a drop in crude prices, etc. Of course, insurers are also exposed to the “objective risk”, that is the relative difference between an insurer's expected losses and their actual losses... but just as an oil company buying an oil deposit is exposed to the relative difference between its expected size (as resulting from geological modelling) and its actual size. The truly specific feature of insurers is to take on the risk *of others*. This core activity is not a question of managing risks *per se*, but of risk pooling – of managing heterogeneity of outcomes among their set of contracts:

an insurer will make use of the fact that, at the end of the year, some of its clients had accidents while others did not. The business model of a financial institution thus rests not on the management of risks⁴⁹ but on the management of heterogeneity (of outcomes). The difference is crucial: asymptotically, when the law of large numbers applies, describing heterogeneity is a predictive matter whereas describing risk, by definition, is not a matter of prediction. As a consequence, “risk models” and related statistics, efficient for pricing and thus managing *heterogeneity*, are not necessarily appropriate tools for *risk* management.

Put another way, the belief that “statistics and actuarial tools are appropriate to manage risk and make decisions when facing uncertainty” is based on the conflation between the points of view of different agents of the risk industry. When signing an insurance contract, for example, the policyholder and the insurer are faced with the same underlying phenomenon – an unknown regarding whether this particular contract will result in compensation – yet the parties are not in the same *situation* with respect to the contract or the unknown: the policyholder, even if he has a certain knowledge of statistics, *does not know* what will happen to himself, whereas the insurer *knows* that, statistically, across all its policyholders, he will pay, approximately, x amount. However, while these two actors are in different situations, we are used to the same formalism being used to describe the common phenomenon with which they are faced when signing the contract, *regardless of their viewpoint*. The distribution function and the statistical indicators such as expectancy or VaR describe both the random phenomenon (here, the outcome seen by the policyholder) and the heterogeneous population (here, the outcomes seen by the insurer). Is this valid?

The answer given to us by physics is: potentially no. Indeed, physics tends to adapt its choice of model to the observer’s position: ice observed over a short

⁴⁹ Which is a crucial support function, fundamental to the operations of insurers, as to many other firms, individuals, organizations or governments.

period of time, for example, to build an ice hotel, will require equations from solid mechanics, whereas the same ice observed over a long period of time, such as for analyzing the flow of a glacier tongue, will require equations from liquid mechanics. By contrast, financial regulations tend to use a single analytic framework, thus disregarding the decision-maker's situation. This can lead to transposing tools for situations where they are relevant onto situations where they are not. It happens for example, regarding prudential questions, when one transposes portfolio assessment solutions on risk management issues. These tools would thus serve as a vehicle for a cognitive representation of phenomena that is not adapted to a correct understanding of the situation, resulting in a weakened quality of regulation and decision-making.

Inspired by Doherty's (2013) idea of endogenizing the observer, this paper positions itself at the level of interaction between the phenomenon and the observer/decision-maker, to identify the situations in which the usual tools and concepts are not relevant for understanding and regulating risks, as opposed to those situations in which they are. The first section will set up the conceptual framework and accordingly characterize the field of relevance of the above-mentioned statistical indicators (and associated estimators), as well as the consequences of their use outside of this field; the reader eager to focus on operational applications can go straightforwardly to the summary ending this part. The second section will focus on the application to prudential regulation and enterprise risk management. It will be organized as follows: first, summarizing the regulatory debate; then presenting our critical analysis; finally proposing solutions.

1. Underlying conceptual framework and derived general results

1.1 Definitions

We focus on characterizing the intersection of an underlying random phenomenon and the decision-maker who seeks to understand it. We define this context in a generic way with the word *situation*:

Definition 1: A situation is a context where a future phenomenon and an agent come together, *such that*:

- (i) *the phenomenon will result in one or more realizations in the eyes of the agent,*
- (ii) *before each realization, the agent does not know what will result from the realization,*
- (iii) *the dispersion of the realizations of which the agent can conceive is significant in regard to his decision-making threshold.*

For example, taking out an insurance policy relates to a unique underlying phenomenon (e.g. whether the policyholder will develop a cancer). The dispersion of the individual outcome is significant, as one would behave in a different way if he knew the outcome before signing the contract. Notwithstanding the uniqueness of the phenomenon, the policyholder's and the insurer's situations are not identical, as the policyholder signs a single contract and observes a single realization, whereas the insurer signs a large number of contracts.

The insurer's situation differs operationally from that of the policyholder in that as he signs a large number of contracts, the underlying phenomenon from his perspective is experienced across multiple individual cases. Taken as a whole, thanks to the law of large numbers, our retail mass-market insurer can forecast (with a certain margin of error of course, as in *any* forecast). The insurer's portfolio will have a foreseeable⁵⁰ technical result. Thus, there is a difference in the nature of the observer's situation depending on whether the underlying phenomenon repeats itself: in one case, the situation remains fully random (significant dispersion of outcomes), whereas in the other it is

⁵⁰ "Deterministic" at the first order, for most retail insurances, where the deviations around forecast are negligible. For some retail insurance (e.g. third-party liability or natural catastrophe) and for most corporate insurances, the deviations may be significant, and the insurer therefore is not in a polar case. Then, managing the heterogeneity of outcomes will also involve managing risk. As a consequence, the insurer sets up risk management actions, e.g. buying a reinsurance treaty.

foreseeable (with business-as-usual margin of error). We can therefore identify two types of situations:

Definition 2: *Theoretically, in an idealistic perspective, random situations are situations where the observer will only observe a single realization; heterogeneous situations are situations where the observer will observe the coexistence of an infinity of realizations (the aggregation of different outcomes, each corresponding to a specific segment of the observed population).*

Operationally, in a pragmatic perspective, we are in a random situation when the number of realizations for the observer in question is “low” and in a heterogeneous situation when the number is “high”. The transition area between “low” and “high” is defined by the consequences held by its crossing: when the eventuality of a significant deviation from an expectation is negligible in the eyes of the decision-maker (i.e. “deterministic” context), it is a heterogeneous situation. In the opposite case (remaining uncertainty non negligible), it is a random situation.

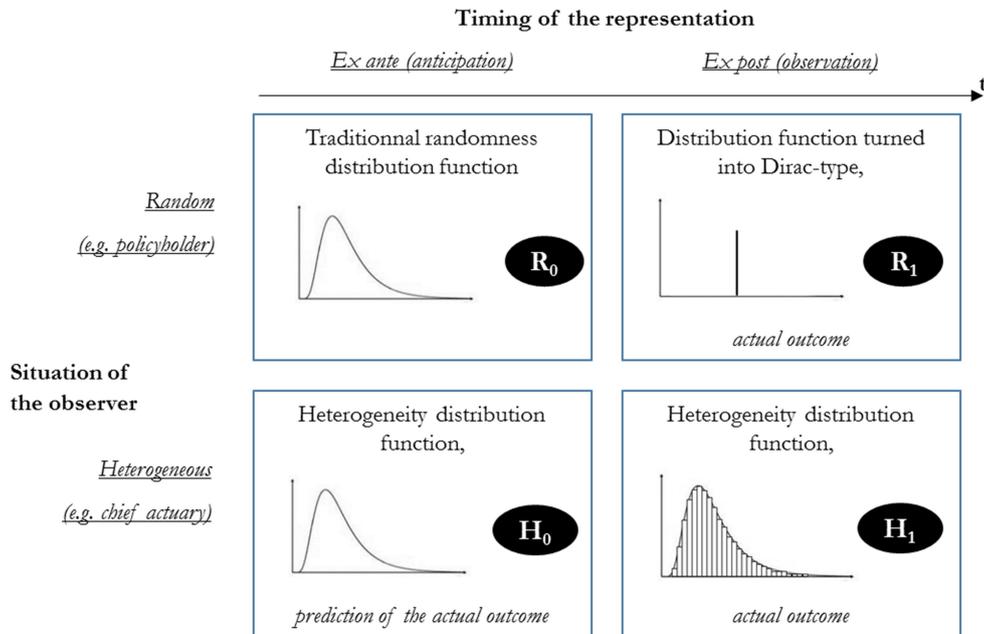
Let us take two examples in an insurance company: the actuary and the Chief Investment Officer (CIO). The actuary prices and sells thousands of similar policies, each of which results (or not) in a compensation, doing so in a *heterogeneous situation*. Conversely, the CIO chooses the strategic asset allocation for a period which is significant in relation to the duration of his tenure, not knowing what will be the overall return of each asset class: he will make this kind of decision few times during his tenure, and as such he is in a *random situation*⁵¹.

⁵¹ In the definition we propose, three points are to be highlighted:

- (i) The definition of the type of a situation is totally *contingent upon the observer*.
- (ii) The distinction between a random situation and a heterogeneous situation is based on the number of *future* realizations of the phenomenon: the precise knowledge of the past does not matter *a priori* in the distinction that we make here between random situations and heterogeneous situations. As a corollary, the distinction between risk and uncertainty (Knight, 1921) is not fundamental in our view. However, one could underscore the *de facto* frequent proximity of random situations and uncertain situations. As a matter of fact, when there is only one realization that occurs in the future (random situation) then, generally, the probability is difficult to assess, either because it is very low and so there are few observations from the past

1.2 Graphic intuition

Let's consider, say, the modeling of the remaining lifetime after the purchase of an annuity. The classic mathematical representations according to the observer's situation are illustrated in figure 1. The first row corresponds to the policyholder who subscribes annuity, living only once (an observer placed in a random situation – R), and the second to the insurer who considers its large portfolio (an observer in a heterogeneous situation – H). The first column corresponds to the moment when they sign the contract – t_0 , before the realization of the outcome; the second column corresponds to the period after the death of the policyholder – t_1 , after the realization of the outcome.



to measure it (e.g. for Solvency II capital requirements, estimates of the 1-year 0.5% VaR of a phenomenon cannot be reliable, see Frezal, 2018), or because the situation is specific and comparison with others to assess the probability is difficult (e.g. what is the probability for a given person to die from a cancer, when this person is unique, and so different from everybody else in terms of way of life, genome, etc.).

(iii) Lastly, in our view, as regards the pertinence of statistical indicators to the basis of a reasoning, the often highlighted opposition between thin-tailed and fat-tailed distributions (e.g. Taleb, 2007) is secondary to the distinction between randomness and heterogeneity. Of course, this affects the number of realizations that mark the threshold between a random situation and a heterogeneous situation, but as we will see, the difference of analysis required to manage these situations has to do with the difference in the nature of these situations and not with the thickness of the distribution tail.

Chart 1: *distribution functions according to situation*

Function R_1 is a Dirac delta function, as the policyholder died only once, a given number of years after signing the contract. Function H_1 is, say, a log-normal function, representing the life length actually observed on the whole set of policyholders. Both R_1 and H_1 represent an *actual* outcome.

Function H_0 is also a log-normal function. Due to the law of large numbers, it properly characterizes the future outcome: it is a deterministic description of the result H_1 with which our observer will be faced. In contrast, the distribution function of R_0 does not correspond to a description of an actual outcome.

The difference in nature between a random situation and a heterogeneous situation appears clearly in the way that the phenomenon's distribution function, in the case of heterogeneity, remains stable in H_0 and H_1 , whereas in the case of randomness it transforms from R_0 to R_1 . In other words, H_0 is a prediction, whereas R_0 is not. The statistical quantities associated with H_0 are estimators that will actually be observed, which is not true of R_0 .

Since *ex post* functions R_1 and H_1 are different after the phenomenon's realization, one might naïvely wonder whether it would be justifiable to assimilate *ex ante* functions R_0 and H_0 , neglecting the observer's situation. These are precisely the points we address: (1) mathematically, function R_0 is not an apt representation of the phenomenon our agent is confronted to in a random situation, and therefore does not provide insights which improve the quality of decision-making. (2) Operationally, such a conflation between these situations generates a biased representation, unfounded expectations and a lack of accountability. It is thus a source of misinterpretation and potentially injurious governance, with negative effect on the quality of decision-making.

1.3 General results

The logical fallacy

Have you ever tried to explain intuitively the meaning of “expected value” without relying on the notion of average? Or a value-at-risk without referring to the quantile that would be observed if the game were to be played several times? Even in a thought experiment, we cannot interpret the concept of, say, expected value otherwise than as embodied in an average⁵² – that is referring not to a single event, but to a whole set of events.

Consequently, in a random situation, all reasoning based on an expected value or any other statistical quantity is based on an implicit condition (repetition) that is impossible to fulfill.

→ **Result 1:** *in a random situation, all reasoning based on an expected value or any other statistical quantity is flawed.*⁵³

Biased apprehension of the situation

When making a decision with the help of a type R_0 function, the intuition we call on to understand the tool we are using rests on the assimilation of type R_0 into type H_0 . As a consequence, in a random situation, the decision-maker using a type R_0 description of the situation is led by this “frequency” metaphor to use a system of thought where the phenomenon repeats, and thus to enter into a mental representation where the situation is deterministic. It gives him the feeling that he *knows* what is going to happen: the use of statistical tools corrupts the observer’s apprehension of the situation, inducing a perception of

⁵² just as, mathematically, expected value can be described only by a sum, be it continuous or discontinuous

⁵³ To explicit it in a logical way, basing a reasoning on statistics while in random situation corresponds to use the following reasoning :

Let A be the future repetition of the realization and let B be the apprehension of the underlying phenomenon and the resulting decision,

*If A, then B
Not A..., then B.*

the risk models as predictive. This can be illustrated by a set of concrete examples, each of them being a direct consequence of this fallacious perception generated by the obliviousness of users as to the change in the nature of their models.

First, the ensuing frequent disappointments with these risk models “*that hadn’t foreseen that...*”. A disappointment that appears also in the academic literature: For example, when it takes over “*criticism of such risk measures [...] because of their poor predictive character*” (Brendel, 2017).

Second, the observation, during fieldwork (participant study), of *several* C-Suite professionals who steer major financial actors and were convinced that their Solvency II risk models were “*predictive*” (and maintained their position when opposed), even though this is contradictory with the very concept of “risk model.”

This second observation may be strengthened by other practitioners who developed a critical reflection on the impact of risk measures on decision making in companies. For example, one explained that, with cat risk measures, “*We create a feeling of confidence. We forget uncertainty*”.

Third, the fact that the expected value may be wrongly perceived as an estimator of the result. Such fallacious perception can be illustrated in the following exchange, observed in fieldwork during a large financial company’s investment committee meeting:

- CFO: “*we have a higher level of risk than the market [than our competitors]: our equity allocation is higher, our real estate allocation is higher...*”
- A critical participant: “*in that case, we should have a higher rate of return. I don’t think that we’re really above it*”.

The expected return – higher than the market standards in the risk/return framework chosen by the company – is seen here as an estimator of what *should* have been obtained. The participant thus considers it abnormal (i.e.,

unnatural, a reason for suspicion, a hint of an error or a lie) that the observed returns are lower.

As a result, the decision makers base their decisions on erroneous mental representations, where the risk has disappeared *de facto* from their field of perception. More broadly, this kind of bias in the representations we make of the world is the foundation from which we make decisions related to randomness on the basis of a cost-effectiveness analysis, such as when signing a reinsurance treaty, a manager compares its price to a deterministic “cost of risk,” or when an investor considers the Sharpe ratio or one of its embodiments. This kind of practice marks the final stage in the process of drifting from the apprehension of risks to a quantified risk/return couple, where risk is formally absorbed into a single synthetic performance indicator. Risk has thus completely disappeared from the decision-maker’s perception of his environment, leaving him with a control panel that has been reduced to one figure and stripped of all randomness.

→ **Result 2:** *By assimilating random situations into heterogeneous situations, the use of statistical indicators leads to a deterministic perception of phenomena and therefore to a distortion of our apprehension of the world, out of which the risk paradoxically tends to disappear.*

Act of faith and absence of responsibility

In a heterogeneous situation (which is to say, when we are faced with function H_0), the situation is deterministic and will transform into function H_1 . It will be possible to measure the gap between the two and assess the modelling error. Whenever it is the case, the *ex ante* description can be proved wrong, which is the characteristic of a scientific approach (Popper, 1959). On the contrary, for an agent in a random situation, the distribution function cannot be back-tested since the agent will always be faced, *ex post*, with a Dirac-type distribution function. Furthermore, having the function R_0 *ex ante* authorizes any function R_1 *ex post*. Therefore there is neither a way to define and measure the potential

modeling error, nor to prove the model wrong. In such a situation, relying on a “risk model” is an act of faith – not science, (and thus induces irresponsibility – not accountability).

Let’s consider for example the fact that seven named hurricanes struck the US during 2004-2005. One could remember that it prompted insurers and cat risk modelers to question the validity of the models that had been used. The probability of one hurricane was low. As a consequence, the probability of seven hurricanes was very low, and it still happened. But who can tell that the probability was false? The mere fact that an extremely low (but not nil) probability event happens is not proof that the estimated probability is false: deciding that the models were to be modified was an act of faith (“I believe that this is a bad model rather than bad luck”), or of social convention between insurers and reinsurers (“We collectively agree to review the modeling and pricing of cat coverage”). Nobody could prove such a low probability, before or after the revision, right or wrong⁵⁴.

→ **Result 3:** *While modelling a heterogeneous situation is scientific and committing, modelling a random situation through a distribution function (or related statistics) is an act of faith with no accountability, as its quality cannot be assessed ex post.*⁵⁵

1.4 Summary of the theoretical framework

We have laid out the theoretical framework that allows us to distinguish two kinds of different phenomena: random situations and heterogeneous situations.

⁵⁴ By contrast, if year after year around ten hurricanes strike the US, then we slowly move into a heterogeneous situation and can assess the reliability of the related model.

⁵⁵ As regards risk models, the notion of *model error* – most often referred to as *model risk* – has been the subject of great interest since the crisis of 2008. Yet the result above can be reformulated as follows: when one observes only one outcome, he cannot distinguish which part of it would be the random deviation that was modelled and which part of it comes from a modelling error. They are fully fungible; as such, risk models – being indistinguishable from their associated model risks – do not exist in a random situation: *while the concept of the dispersion model makes sense in a heterogeneous situation, the very concept of a risk model is not operationally pertinent in a random situation.*

- (1) When a decision-maker is confronted with a random event that will occur only *once in the context of his decision-making*, he is in a *random situation*. If, *within the context of his decision-making*, this event occurs an *infinite* number of times, he is in a *heterogeneous situation*.
- (2) In heterogeneous situations, the decision-maker has an almost deterministic understanding of the world: he knows *ex ante* the distribution function of the events he will observe *ex post*. This is not the case in random situations.
- (3) This alternatively predictable or unpredictable character of the result is what allows us, more generally, to demarcate heterogeneous phenomena from random phenomena in non-polar cases where the number of future realizations is more than 1 but is not infinite: *if the forecast error for the distribution function observed ex post by the decision-maker is negligible in his eyes, it is a heterogeneous situation. If not, it is a random situation.*

In random situations, the reliance on statistical quantities is based on a conflation of randomness and heterogeneity. The principle consequences of this are as follows:

In a random situation,

- (5) To base a decision on the statistical quantities associated with the distribution function expected *ex ante* relies on a fallacious reasoning. Such an analysis alters the decision-maker's perception of the phenomenon, giving him an illusion of determinism, skewing his reasoning and impairing his ability to take an appropriate decision;
- (6) Risk model and model risk become indistinguishable; it is impossible to determine accountability regarding the model's quality *ex post*.

Legal heuristics tend to demonstrate that our societies are used to differentiate between situations of risks and situation of heterogeneity (see appendix A). However, some prudential regulations tend to conflate the two. The following section aim to clarify such misunderstandings and suggest superior bases.

2. Capital requirements

2.1 The regulatory debate

The last evolutions of the main prudential standards revealed distinct choices. The U.S risk-based-capital (RBC) system built in the 1990's and Solvency I (which was consolidated in Europe until the early 2000's), although taking into account a hierarchy of risks,⁵⁶ do not pretend to determine the capital requirements as a risk measure of the own funds.

By contrast, for the past fifteen years, the EU has been implementing a new prudential regulation in the insurance sector (European Commission, 2009) that, in its quantitative dimension, pretends its capital requirements to correspond to *the* annual VaR of 1/200 of the own funds. This regulation is often perceived, or presented by its promoters, as cutting-edge. However, relying thus on a statistical measure to assess the solvency of a company – that is, for strategic risk management purposes, dealing with the forthcoming year survival or bankruptcy of a company, is basing the regulation on a conflation between heterogeneity (statistics on the whole industry or over a long period) and randomness (dead or alive at the end of the year). As a consequence, the set of general results which we established in Part I applies and we will successively present the operational embodiment of results (i), (ii) and (iii).

2.2 The critical analysis

The interpretation and calibration issue

The regulation fixes the calibration of the capital requirements to the 99.5% VaR, which seems to be mathematically univocal. But how should we interpret that? As always, interpreting a probability measure requires positing oneself in a heterogeneous situation: some say that with such a regulation, no more than one company out of 200 will collapse each year; others say that a given company will be able to face a stress with a return period of 200 years.

⁵⁶ Solvency I does so in a very rustic way, only distinguishing UL from minimum guaranteed rate contracts in Life and liability insurance from other lines of business in non-life. The American RBC standards are based on finer characterization of the risk, including per asset types.

These two interpretations are not equivalent as they generate different calibrations (Planchet and Leroy, 2010). If we expect approximately one of every 200 companies to go bankrupt each year, then we should look at the dispersion of the different companies that make up the industry and review the calibrations annually. If we expect that each company statistically goes bankrupt one year out of every 200 years or so, then we should look instead at the dispersion of the situation in time and the calibration, company by company, should remain stable over time. In terms of objective, the first option would relate to idiosyncratic risk and correspond to “*the traditional solvency risk controlled by the existing solvency regulation that are aimed at counterweighting the moral hazard problem*” (Bobtcheff et al., 2016) whereas the second option corresponds to a kind of systemic risk management, and interacts with the procyclicality issue (Frezal et al., 2016). In terms of concrete calibration, neither option makes sense: the first option requires us to measure the worst deviation between 200 firms, and considering it for dissimilar companies would be meaningless; the second option leads us to measure the worst deviation over two centuries, a period covering so disparate years.

Mathematics of randomness and of heterogeneity are the same, so the regulatory quantum seemed univocally defined, but since randomness and heterogeneity are distinct issues (a distinction unaddressed by the regulators), the targeted scope of heterogeneity has never been defined: no one knows whether the operational interpretation and calibration of the quantile should be temporal (option 2) or “geographical” (option 1). The objective is unclear, the interpretation equivocal and the calibration undefined. As a result, the European regulator used an inconsistent mix of the two options: it calibrated the non-life underwriting components of the standard formula following the

first option (CEIOPS, 2010)⁵⁷ and the market risks following the second option:⁵⁸ regulation is not technically sound.

A regulation distorting apprehension

From a general viewpoint, the fact that banks and their regulators tended to rely excessively on quantitative models before the 2008 crisis (Ashby, 2011) can be explained by the fact that they were accustomed to using successful models in a business/pricing context. When these models were transposed onto a steering/risks context, the nature of the models changed – they transformed from predictive models (heterogeneity) to non-predictive models (randomness) – and this change in their nature has neither been recognized by their users nor by their environment.

Prudential measures set up by Solvency II, as risk measures, are, by definition, not predictions and thus may generate such kind of serious issues. The observed evolution of own funds at the end of the year will be higher (often) or lower (hopefully seldom) than the capital requirement. However, as already mentioned, fieldworks revealed managers having the view that their prudential models were predictive. Believing the figures one uses are accurate predictions of the future means one considers the world to be deterministic, and thus is not well equipped cognitively to judge the situation.

Indeed, the use of such tools by experienced professional thus positions them in a fallacious state of mind; the risk tends to disappear from their unconscious steering dashboard: decision making under uncertainty appears no more as a matter of strategic choice but only as the mechanistic optimization of a ratio: New Business Value divided by Solvency Capital Requirement (SCR) to decide

⁵⁷ The methodologies used to determine the stresses for premium and reserving risks, which depend on the lines of business, may include methods which take into account the dispersion of combined ratios or reserving result between the different companies on the European market.

⁵⁸ The calibrations are based on the financial markets variability in time and are not revised annually, apart from the equity stress parameter which was decreased in QIS 5 from 39% to 30% to mitigate procyclicality.

whether or not to launch a product, expected result divided by gain in SCR to decide whether or not to enter a reinsurance program, expected return divided by increase in SCR to decide whether or not to invest in riskier assets, etc. When calculating risk in this manner, the consciousness of consequences, which is so decisive in the risk return trade-off, disappears, only to be replaced by a new definition of the return – a “risk adjusted [expected] return” which will concretely never be observed.

Towards the end of accountability

In her paper on risk management in the banking industry, Mikes (2011) proposes a classification of Chief Risk Officers (CROs) into two categories: “quantitative enthusiasts” and “quantitative sceptics”. She explains that while the quantitative sceptics tend to seek impact on management choices and feel responsible if things go wrong, the quantitative enthusiasts rather consider that they did their job providing a risk quantification and that they are in no way liable for the decisions taken.

This can be explained by our framework, as prudential regulation asks CROs to deal with major potential adverse events endangering the survival of the company: regarding these events, top management is in a random situation. First, Mikes’ observation is thus a direct and natural consequence of our general observation of the fact that, in a random situation, no risk model can be proved wrong: when positioning himself as the top expert of risk measurement, the CRO places himself in a situation where it becomes impossible, *ex post*, to reproach him with having made a technical mistake⁵⁹. Second, it may be the consequence of the fact that, in a random situation, no adverse outcome, even extreme, is formally excluded by any risk measure – it is only estimated as highly improbable: even if positioning himself as a decision-

⁵⁹ This can be observed, in particular, as regulation authorizes companies to use internal models to determine the 1/200 VaR while the company is in a random situation in regard to its ability to survive. Due to the porosity between the risk model and the model risk and the resulting lack of accountability in such a situation, one cannot scientifically objectivize *ex ante* nor *ex post* the adequacy of those internal models.

making participant, the quantitative enthusiast CRO acts *de facto* as a formalizer of “bad luck,” offering *ex post* a justification of the decision as “optimal, given what was scientifically known when the decision was taken, even if it proved to be catastrophic afterwards.”

Furthermore, our observations during fieldwork in a major insurance company extends this propensity of risk measures to make people no longer accountable for their actions to the managers in the business units. For example, implementation of Solvency II acted as an argument for operational decision makers to no longer care about risks: when the risk department representatives challenged them on the risk dimension of a new product launch or an investment decision, their reaction regularly was: “*you [risk department] calculated risks and we [the company] have capital to face it, so why, on top of that, do you ask me to care about it? Do you mean that your risk modelling is inappropriate?*”

2.3 The operational solution

The problem to be solved

The main problems arising from the use of statistical quantities to apprehend randomness when regulating a company are the following:

- (i) A regulation as Solvency II is grounded on quantitative foundations that are both equivocal and technically complex, where the underlying choices do not correspond to a concrete interpretation which the political level could grasp. In other words, while the calibration of the required resistance is supposed to be the result of an assumed political trade-off with issues other than prudential, the measure, being not interpretable, has been *de facto* delegated to the technical analysts, so that the decision has not been made at the right level.
- (ii) As soon as the risk is purportedly quantified, and subsumed in one figure, it is no longer perceived as a risk. Thus, the Solvency II indicators, which are used to steer the companies, skew the

decisions by distorting the decision-maker's perception of the situation.

- (iii) Moreover, given that a risk-taking decision-maker is considered rational if and only if he uses the results from such types of calibrations, and that the quality of these comprehensive statistical calibrations provided by the analysts can never be evaluated *ex post*, it is no longer possible to establish accountability for the quality of the decision-making.

In order to avoid these flaws, the tools with which to regulate (capital requirements) and base an internal decision (risk appetite framework) under uncertainty must allow for *the feeling of randomness to be reintroduced, so that the perception of the random nature of the situation remains* and for *the decision not to be driven by uninterpretable analysts' subjective choices, so that it is taken at the right level*.

Towards a solution

A structured scenario-based approach allows answering these issues through a two-layered system:

- The first layer would allow determining the capital requirements and providing an answer to point (i). It would define a set of scenarios (financial, underwriting and others, eventually intertwined) at the regulatory level, devoid of reference to any VaR indicator. Capital requirements would correspond to the amount of capital needed to survive any of these scenarios. The choice of this set of scenarios by the regulatory authorities would allow consideration of the political trade-off between prudential issues and other socio-economic and macro-financial issues.
- The second layer would be part of the company's risk appetite framework and provide an answer to points (ii) and (iii). It would define a set of *ad hoc* scenarios at the company level, by the top

management, and check that they would not generate bankruptcy. These scenarios would be defined as the scenarios (1) which occurrence the top management would consider not negligible and (2) under which the company should, *given its risk appetite*, resist.⁶⁰

Two points are here to be highlighted. First, in this framework, the strategic risk management decision, both at the regulatory level and at the company level, resides in the definition of the set of scenarios. This set respectively defines the risk appetite of the society (regulatory level, capital requirements) and of the company (top management level, risk appetite framework). Should the company be able to resist a home country government bonds default, or do we accept bankruptcy in the case of such event? Should it resist a Great-Depression-like financial crisis? A Spanish-flu-like epidemic? A combination of these two last events?

Second, this approach does not exclude quantitative modeling, as assessing the impact of scenarios on the balance sheet (and/or the result) requires strong technical capabilities. This combined approach only leads to restoration of the strategic risk management decisions (which circumstances shall our company be able to resist? Which risks are we ready to take?) in advance of the quantitative technical filter, so that it reduces the impact of the experts' theorization (e.g. choice of a silo-based approach) and subjectivity (calibration choices).

Pros and Cons

Renouncing statistics requires that we renounce (i) the ambition to compact the potential outcomes to the point of reducing them to one or several deterministic statistical indicators that claim to synthesize all possibilities and

⁶⁰ Such a double-layered approach is compatible with the current regulatory trends towards the incorporation of a qualitative dimension. Our second layer is a kind of *Major Risks Self Assessment* exercise, as that which exists in Pillar 2 (qualitative requirements) of Solvency II. US regulators also contend that an overemphasis on capital requirements is problematic and that other regulations and methods should be employed in the supervision of insurers' financial condition and risk (e.g. early warning systems, financial exams, etc.).

(ii) the ambition to manage risks through a purely technical analysis, with a dichotomy between risk calibration experts and decision-makers. Furthermore, from a psychological viewpoint, this process is harder to bear for the decision-maker insofar as he must take responsibility for the discretionary aspect of his decision – defining, as a negative, the scenarios under which he accepts bankruptcy, and cannot shift blame onto a hidden technician or an ethereal statistic.

However, such a process is intuitive as it relies on an instinctive decision-making process. As a matter of fact, according to Solé (2000), individuals make their decisions by obliviously categorizing future scenarios into three classes: that which cannot occur (Impossibles), that which could occur (Possibles), and that which is inevitable (Non-Impossibles, which we cannot imagine not occurring). These mental representations are largely the unconscious result of each individual's culture and experience. They define our perception of reality: a mental schema of analysis that frames decision-making for each of us.⁶¹ Our proposal is nothing but a formalized – explicit and thus generating awareness and accountability – process of this oblivious decision-making schema: the set of scenarios which are to be determined by the regulation and the top management reflect their strategic view on the worst “Possibles” to which the company should resist.

⁶¹ Solé takes the example of the Pearl Harbour attack: at a dawn of December 1941, the radar operators in Pearl Harbor observing points on their screens saw a dysfunction of this recent technology as a Possible, and an attack without declaration of war as an Impossible; such a representation of the future led them to spoil two hours checking their radar rather than alerting and drifting the ships out of the harbor. This representation of reality defines the strategy and is therefore necessarily subjective (obviously, the Japanese pilots, at the same moment, did not have the same representation of the future) and, as a consequence, is the prerogative of these who are legitimate to arbitrate trade-offs: top-managers and politicians. In the Appendix B, we propose a mathematically formalized description of this decision making process.

Our approach corresponds to a focus on the possible consequences of our decisions, which also corresponds to Taleb's recommendations (2007).⁶² As such, in terms of asset allocation, a 1/N allocation strategy such as Taleb (op. cit.) or Haldane and Madouros (2012) recommend corresponds to not wanting to exclude from the field of Possibles any scenario in which a given asset's value falls. The old rules of diversifying investments to which European insurance companies cleaved in Solvency I, and which limited the proportion of the balance sheet invested by issuer category and by issuer, are an embodiment of this. It also corresponds to the popular wisdom that has filtered into the proverb "don't put all your eggs in one basket" in several languages, and not "put all your eggs in a quilted basket" (toward which risk-based but ban-free regulations, such as Solvency II, tend).

Conclusion

"[To make] a strenuous effort to put aside some of the artificial schema we interpose unknowingly between reality and us. What is required is that we should break with certain habits of thinking and perceiving that have become natural to us. We must return to the direct perception"

H. Bergson, *La Pensée et le mouvant*, 1938

(Second lecture at the University of Oxford, 1911)

Tools that may damage the quality of decision-making

The economic and financial tools for apprehending risks on the basis of statistical indicators are founded on the conflation of two phenomena of different natures: heterogeneity, for which statistical concepts have emerged and in reference to which we continue to intuit these concepts, and randomness, onto which these concepts have been wrongly transposed.

⁶² but for different reasons. Taleb considers that we cannot know the probabilities of rare events; we judge that as soon as we decide to consider their occurrence as possible, their probability is not an issue.

The pertinence of certain tools is contingent on the observer's position in respect to the phenomenon he seeks to understand. The transposition onto random situations of tools conceived for understanding heterogeneity is exceedingly easy from the standpoint of mathematical formalism, but it is based on the identification of two decision-makers whose positions are radically different. If, for observers facing heterogeneity, statistical methods provide meaningful insights into decision-making, for observers in a random situation, this transposition, which leads to the use of statistical indicators,

- (i) is not pertinent because it relies on a logical mis-reasoning,
- (ii) gives rise to an erroneous mental representation of the situation by creating an illusion of predictability,
- (iii) and does not provide the means to make the actors aware of their responsibility.

Hence, the “information” provided is likely to harm the quality of the decision-making.

Regulatory consequences

Common visions over prudential regulations (together with quantitative risk management practices) are based on the focus or optimization of statistical indicators (expected value, variance, VaR, etc.) and are therefore subject to these failings.

Hence, our analysis dismisses the Solvency-II-Pillar-1-like capital requirements as a relevant regulatory framework. Indeed, the reference to the 99.5% VaR indicator blurs the real issues behind a technical, and – as we have tried to demonstrate – fallacious approach. It displaces the debate to a chimeric statistical field which is in fact neither more objective nor more legitimate. This advocates for other procedures to be employed by regulators such as stress test-like approaches. Thus, a set of common scenarios would be defined by the public authority (e.g. a 1929 financial crisis; the combination of a Andrew-type

hurricane with Hong-Kong flu-like epidemic and a 1998 Russian financial crisis; a long-term stagflation period combined with a rise in longevity; etc.), and the capital requirements would correspond to the amount needed to keep the own funds positive under each of these scenarios. The choice of such set of scenarios should invoke an accepted responsibility on risk-taking (the risk appetite of the society we live in) and the legitimate arbitrage between the protection of the policyholders and the multi-dimensional impacts of any financial regulation in the economy. It is in accordance with the emphasized importance of ORSA to assess the risk profile of a company (Fischer and Schlütter, 2015), this time at the macroeconomic level. It could be easily combined with the identification of sets of scenarios peculiar to each firm, defined by its top-management and making its specific risk appetite explicit. The crucial point here is that the *actual* decision is made when defining the set of scenarios to which the companies should resist. As a consequence, this methodology not only improves the pertinence of the decision-making framework, but also prohibits actors from abdicating responsibility.

Fields of impact to explore further

From a theoretical standpoint, the distinction between randomness and heterogeneity provides a rationale to explain some legal heuristics, such as the apparent paradox of the coexistence of prudential regulations, which aim to prevent financial institutions from taking as many risks as they like, and of limited liability companies, which on the contrary encourage them and their controlling shareholders to take more risks by protecting them from being held responsible for unlimited losses (see appendix A).

From an accounting standpoint, this distinction leads to question the relevance of stochastic Present Value of Future Profits (PVFP), stochastic Best Estimate of Liabilities (BEL) and Time Value of Option and Guarantees (TVOG) when valuing a life insurance company balance sheet. Indeed, while an insurer faces the heterogeneity of numerous contracts on the underwriting side, he is in

random situation regarding the financial markets evolution. As a consequence, while considering the expectancy to set the reserve is relevant with respect to underwriting parameters, considering expectancies among different financial scenarios rather than a central scenario, as it is done when calculating a TVOG, could appear meaningless.

From an operational standpoint, the fruitfulness of this distinction in the field of risk management must be explored further. Regarding financial regulation, for example, it can suggest evolutions in governance. Indeed, activities concerning the calculation of capital requirements are generally attributed to the CRO, due to his responsibility for “capturing” risks. However, this calibration rests on a foundation shared with profitability measurement tools and is arbitrary from the viewpoint of decisive risks. As such, this would seem to argue in favor of such responsibilities being steered by the Chief Financial Officer along with the other balance sheet parameters. Conversely, as risk management necessitates the imagination of possible futures, it would make sense to entrust strategy function to CROs.

References

- Ashby, S., 2011. “Risk Management and the Global Banking Crisis: Lessons for Insurance Solvency Regulation”, *The Geneva Papers on Risk and Insurance – Issues and Practice*, 36, pp. 330-347.
- Bergson, H., 1938. *La Pensée et le mouvant*, Paris, Presses Universitaires de France. (Ch. V, Conférences faites à l’Université d’Oxford les 26 et 27 mai 1911).
- Bobtcheff, C., Chaney, T. and Gollier, C., 2016. “Analysis of Systemic Risk in the Insurance Industry”, *Geneva Risk and Insurance Review*, 41 (1), pp. 73-106.
- Brendel, C., 2017. “Systemic Uncertainty and the regulation of complex financial systems”, *seminar paper*.

Committee of European Insurance and Occupational Pensions Supervisors, 2012. *CEIOPS' Advice for Level 2 Implementing Measures on Solvency II: SCR Standard Formula Calibration of Non-life Underwriting Risk*, DOC-67-10, eiopa.europa.eu.

Doherty, N., 2013. "Risk and the Endogenous Economist: Some Comparisons of the Treatment of Risk in Physics and Economics", *Geneva Risk and Insurance Review*, 38 (1), pp. 1-22.

European Commission, 2009. Directive 2009/138/EC of the European Parliament and of the Council on 25 November 2009 on the taking-up and pursuit of the business of Insurance and Reinsurance (Solvency II), *Official Journal of the European Union*, L335/1, 17.12.09.

Fischer, K. and Schlütter, S., 2015. "Optimal Investment Strategies for Insurance Companies when Capital Requirements are Imposed by a Standard Formula", *Geneva Risk and Insurance Review*, 40 (1), pp. 15-40.

Frezal S., Haguët, E. and Nou, V., 2016. "Procyclicité : inéluctabilité, avantages et inconvénients", *Working paper PARI*, <http://www.chaire-pari.fr/publications/#working-papers>.

Frezal S., 2018. "Solvency II Is Not Risk-Based—Could it Be? Evidence from Non-Life Calibrations", *North American Actuarial Journal*, 22(1), 15p.

Haldane A. and Madouros V., 2012. "The Dog and the Frisbee", speech, *Bank of England*.

Knight F., 1921. *Risk, Uncertainty and Profit*, Boston and New York, Houghton Mifflin.

Mikes, A., 2011. "From Counting Risk to Making Risk Count: Boundary-Work in Risk Management", *Accounting, Organizations and Society*, 36 (4), pp. 226-245.

Planchet, F. and Leroy, G., 2010. "Que signifie la ruine dans Solvabilité 2", *La Tribune de l'Assurance*, 147, pp. 55-56.

Popper, K., 1959. *The Logic of Scientific Discovery*, Hutchinson & Co.

Solé, A., 2000. *Créateurs de mondes, nos possibles, nos impossibles*, Monaco, Editions du Rocher.

Taleb, N. N., 2007. *The Black Swan*, New York, Random House.

Taleb, N. N., Read R., Douady R., Norman J., and Bar-Yam Y., 2014. “The Precautionary Principle (with Application to the Genetic Modification of Organisms)”, *NYU School of Engineering Working Papers Series*.

Appendix

Appendix A: Legal heuristics which tackle this conflation

As a risk that is too high for an individual can thus be constitutive of a heterogeneity that is beneficial to society and vice versa, this distinction helps to explain certain juridical mechanisms as heuristics developed by our societies to lead individual decision-making through mechanisms which alternatively curb or promote risk taking (i.e. incentivize desirable risk behaviors). In the first category, we find examples such as prudential regulations, which aim to prevent financial institutions and their executives from taking as many risks as they like so as to limit bankruptcy. In contrast, the second category includes examples such as limited liability companies, which protect executives and shareholders from being held responsible for unlimited losses and protect them from the claims of creditors in the case of bankruptcy. What is the difference between these two categories? In the first case, it is the savers who suffer the consequences of bankruptcy, and in the second, it is the suppliers. Savers generally do not open several retirement or savings accounts – and the financial system is exposed to systemic risk: they are in a random situation regarding the possible bankruptcy of their bankers or life insurers. The issue is thus no longer expected value: the possibility of ruin would be too painful for the savers, so it is necessary to limit the potential and the impact of such an eventuality. By limiting risk taking that would be optimal for the shareholders of the regulated company (and who diversify their investments, thereby placing

themselves in a heterogeneous situation that may push for each company to take risks), society protects the savers, who are in a random situation. Conversely, if a standard non-financial company goes bankrupt, it is said company's suppliers who will not be reimbursed and will lose a part of their revenue. Most suppliers, however, are in a heterogeneous situation: statistically, they already know they will suffer losses due to one client or another, and a gains/losses balance can then be established socially. Such a balance is effective and must not be disturbed by an executive, who would be in a random situation if he had to make up for the bankruptcy with his own assets. Society may limit this potential disturbance of social efficiency by limiting the small business executive's responsibility. Such comparisons raise the question of whether it is appropriate to maintain a limited liability for shareholders and top managers of financial institutions.

Appendix B: Mathematical formalization of an *à la Solé* decision making process under uncertainty

Let us propose an intuition of the process of analysis and decision, which is based on the following steps:

- (i) Identification of all possible states of nature: Opening the decision-maker's field of potential Possibles.

In the case of an asset allocation decision, for example, a typology of possible scenarios could be: (1) rise of the spread of government bonds accompanied by a drop in stock prices and a rise in property prices; (2) negative rates accompanied by a rise in stock and property prices; (3) etc.

- (ii) After the most exhaustive inventory possible: The rejection of those scenarios that would be considered "Impossible" *by the decision-maker* so as to retain only *his* "Possibles".

By "Impossibles," we mean those scenarios for which the decision-maker does not want to take into account the Possible realization, either because he judges it to be negligible (literally Solé's

Impossible) or because he voluntarily assumes the risk. In the case of an asset allocation, for example, the decision-maker would have to take a position on whether he considers a default of his country's government bonds as an "Impossible."

- (iii) The analysis of the impact of various possible actions in each of the remaining "Possible" scenarios: The choice of the action that maximizes the decision-maker's utility in the scenario where he "believes" most, under the condition that the consequences are acceptable in each of the Possible cases. In our investment example, this would consist in choosing the allocation that maximizes the returns in the preferred scenario, under the condition that these returns remain acceptable in his other Possible scenarios.

Such a process may surprise because it does not make use of probabilities as such, only the parameter "negligible or not;" but from a normative perspective, it is consistent with the lack of meaning of statistical indicators (such as probabilities) in a random situation, and from a descriptive viewpoint, it seems a natural fit with Solé's analyses.⁶³

⁶³ Of course, this contradicts Arrow's apologue (1951), which uses the following example to justify the claim that a decision-making reasoning under risk cannot do without probabilities: "If an individual were told to predict whether or not two heads would come up in successive throws of a fair coin and further informed that he would lose his life if he guessed wrong, I find it very hard to believe that he would disregard the evidence of the calculus of probability[...] [A]n extension of this suggests that in almost any reasonable view of probability theory the probability of a single event must still be the basis of action where there are genuine probabilities." But, beyond the fact that he ignores the central problem of reference class for a choice made in a more realistic context, this example overly simplifies another aspect of "real life". A choice such as the one proposed by Arrow is between $-\infty$ (to die) with a probability p ($1/4$) or with probability $1-p$ ($3/4$). In a real situation, even a highly simplified one, an individual is not faced with the choice between A with probability p and A with probability $1-p$. Let us suppose for example that for your retirement, you can put your 1 M\$ savings into a lottery that will pay out 2 M\$ if you win. This means that you have the choice between ultimately holding onto 1 M\$ with the probability $p=1$ or playing for 2 M\$ with the probability p' vs. 0 with probability $1-p'$. If we ask the question "does the result depend on whether p' equals $1/4$ or $3/4$?", Arrow's claim becomes much less self-evident. The only issue at stake here, in deciding whether or not to make the bet, becomes: "can I ignore the

Let us propose a formalized version of this process:

*Let \mathcal{S} be the set of all existing states of nature,
of which \mathcal{S}_i is the set of all states of nature that agent i imagines,
of which $\bar{\mathcal{S}}_i$ is the set of all states of nature that agent i identifies as Possibles.*

For example, the case of a giant meteorite razing Paris $\in \{\mathcal{S} - \mathcal{S}_i\}$; the case of a centennial Paris flood recurring this year $\in \bar{\mathcal{S}}_i$ for the homeland security (which is prepared for such an eventuality), but $\in \{\mathcal{S}_i - \bar{\mathcal{S}}_i\}$ for the organizers of Roland-Garros, who certainly accept to disregard the possibility of such an occurrence, or, in other words, run the risk of being unprepared for a flood (and for which no one would blame them).

possibility of losing my savings?” or, in other words, “does losing belong to what I subjectively consider an Impossible?”.

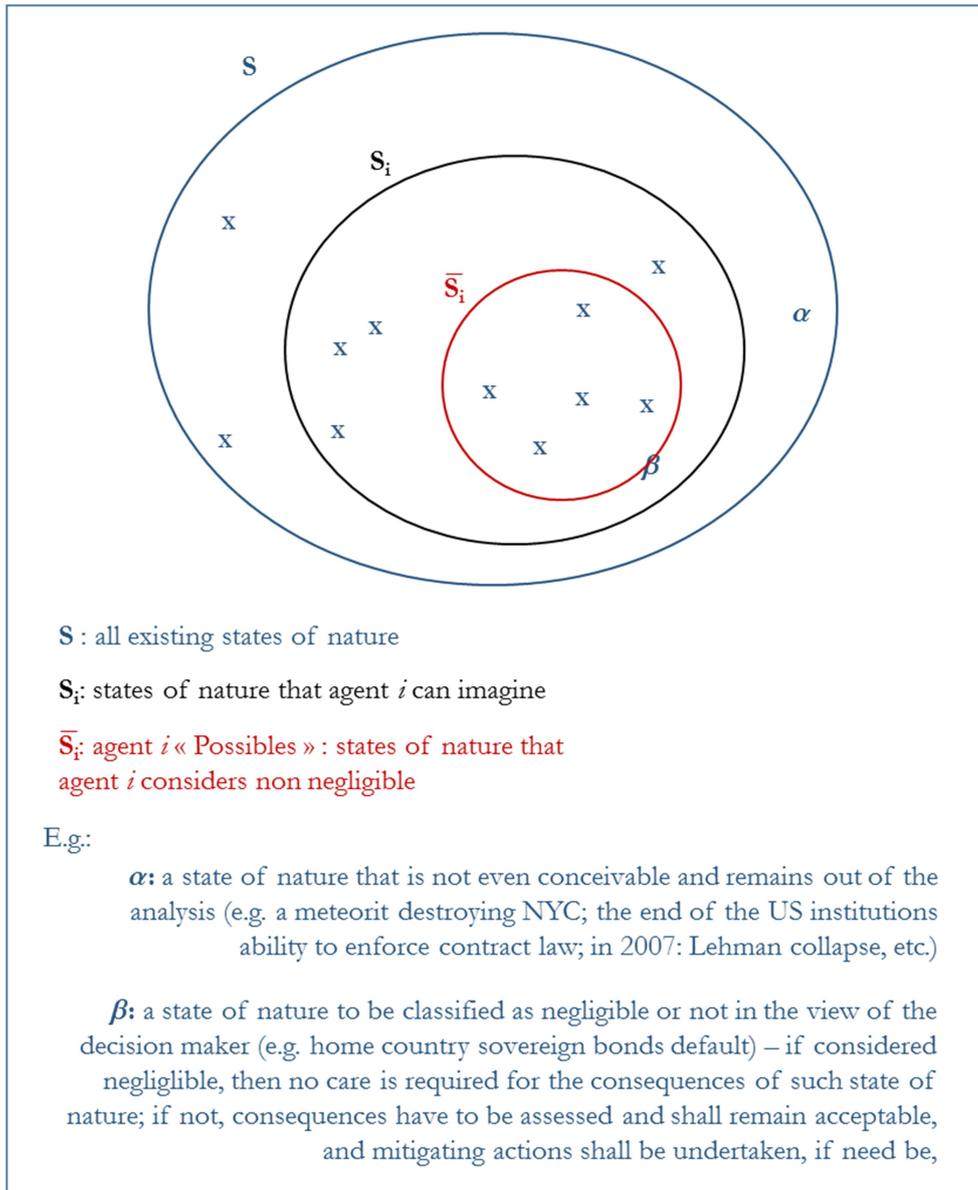


Illustration 1: Preliminary analysis – the field of Possibles and its representation for agent i

Step 1: Widen **S_i** to the most – purely analytical (led by experts)

Step 2: Define **S_i-bar** – analysis framing the forthcoming decision (led by decision makers): *de facto*, it represents a definition of the risk appetite

*Likewise, let **A** be the set of actions a human being can imagine,*

of which A_i is the set of actions agent i imagines,

of which \bar{A}_i is the set of actions agent i identifies as Possibles.

For example, the action of killing one's parents to inherit their wealth faster certainly $\in \{A - A_i\}$ because the idea would not even occur to us; according to the person, the case of cheating on one's partner belongs either to \bar{A}_i , or to $\{A_i - \bar{A}_i\}$ if, as a matter of principle, i considers it taboo; if it belongs to \bar{A}_i , then the question of cheating – which is to say the consequences of the risk – will arise.

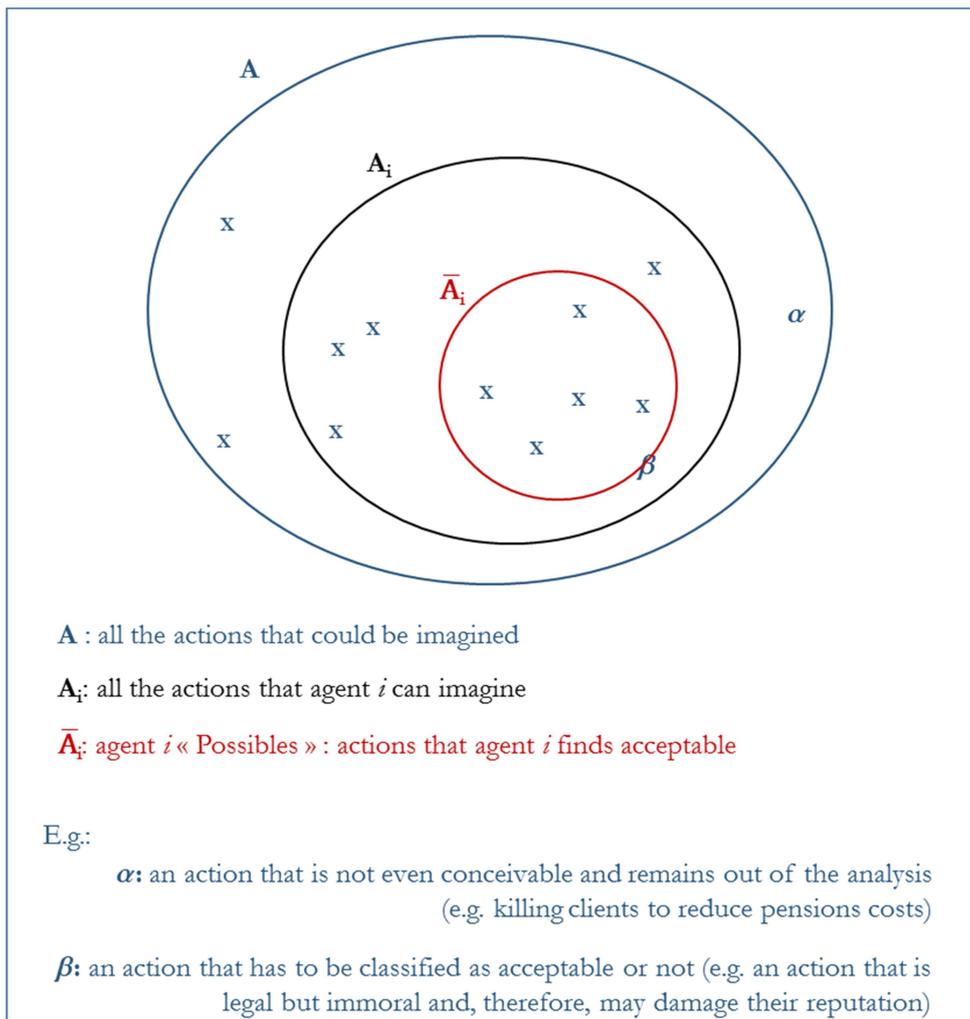


Illustration 2: Agent I path towards decision

Step 1^{bis}: Identify A_i (if need be, thinking out of the box. For example, when facing excessive expenses, improve service to get a premium position rather than reducing costs)

Step 2^{bis}: Define \bar{A}_i (values definition, reality principles, etc.)

Now, let \dot{S}_i be the state of nature in which agent i believes most

Let us note that $U(a; s)$, the pay-off associated with the action a , combined with the state of nature s , and \bar{U}_i , the worst result that i finds acceptable.

The decision making process of agent I then becomes:

- (1) Maximising S_i and A_i (getting the greatest set)*
- (2) Defining \bar{S}_i and \bar{A}_i , \dot{S}_i and \bar{U}_i ,*
- (3) Mere optimization: retaining the decision \dot{A}_i such that its pay-off \dot{U}_i satisfies:*

$$\dot{U}_i = \begin{cases} \max_{A_i} (U(a; \dot{S}_i)) \\ \text{under constraint } \min_{\bar{S}_i} (U(\dot{A}_i, s)) \geq \bar{U}_i \end{cases}$$

Let us note here that it is possible to draw a certain number of parallels with some usual concepts of decision theory: the fact that A_i and S_i are only subsets of A and S corresponds to a Simon-type bounded rationality; the subjectivity associated with \bar{A}_i , \bar{S}_i , and \dot{S}_i is linked to the fact that we consider subjective probabilities in expected utility; and finally, \bar{U}_i can be understood as an equivalent to a parameter such as risk aversion.

There is, however, a fundamental difference between the two approaches. The *decision* here does not reside in the mechanical operation of constrained optimization (step 3) as it does implicitly in decision theory. *What we highlight is the fact that the analysis is (the most important) part of the decision: the decision resides in the choice of \bar{A}_i , \bar{S}_i , \dot{S}_i and \bar{U}_i – a choice that is committed and conscious, and not a passive characteristic of the individual.*

This choice is irreducibly a-rational (and subjective). Indeed, when one does not know what will happen, there is no good decision *in itself*.

Chapitre 4 : Coût et tarification des risques : de Solvabilité 2 à la question de l'équité des prix des assurances

Fairness in Uncertainty: some Limits and Misinterpretations of Actuarial Fairness

Introduction

With the rise of social networks, the arrival of connected objects, the development of information technology, and advances in the data sciences, big data enables increasingly finer personalization of products. In domains other than insurance, studies already show that online prices are adjusted to customers according to the way they are tagged by marketing agencies, rather than the cost of production (Turow, McGuigan, and Maris 2015). In insurance, the possibility to adjust the premium to the individual risk, also called “actuarial fairness,” is empowered by these new technologies, as a vast array of new parameters are made available for a fine tuning of risk measurement and pricing. Some of the data at stake do not obviously affect risks though, but characterize the individual in his propensity to buy or cancel a contract (Siegel 2016). The emergence of data science revives and enlarges, we would like to argue, old debates around the fairness of insurance prices that have never been completely settled.

In this endeavor, insurance regulators have a specific role to play. Yet driven by the same desire to avoid unfair discrimination and trying to put limits on the legitimate parameters, European and American regulators have apparently reached opposite conclusions. For example, albeit women have a higher life expectancy and statistically drive more safely, European regulation has recently prohibited ratings based on sex, “in order to ensure equal treatment between men and women” (European Council, 2004; confirmed by the Court of Justice of the European Union, 2011). In this view, fairness, defined as “equal treatment,” disregards statistical arguments and thus leads to put limits to the adjusting of prices to the cost of risk.

By contrast, the Consumer Federation of America (CFA, 2013), in his criticism of price optimization and the fixing of margins based on the willingness to pay, stresses that “moving pricing away from its historic cost-based approach (...) <is> unfairly discriminatory.” This position follows a long tradition that judges the expected loss, or the pure premium, to be the frame of reference for

determining the fairness of insurance premiums (Daston 1988, 1989; Feller, 1968, 249). In this strand of thought, American insurance commissioners have traditionally given as a guideline that rates should be adequate – i.e. sufficient to cover the costs and maintain solvency, yet reasonable and “not unfairly discriminatory.” Unfairly discriminatory rates were defined for the first time in Kansas in 1909 as “charging different rates to persons with risks of a like kind and hazard” (Miller 2009, 278). This definition is confirmed to this day in similar terms by the Insurance Commissioner of California, who states that “a fundamental tenet of actuarial standards, [is] that rates be based on the expected value of all future costs” (Jones, 2013).

The European Union thus seems to discard some cost-based approaches, whereas the United States reaffirm that it is the only acceptable one, both in an effort to be fair: put in other words, the European endeavor to ensure equal treatment of men and women stated in the “anti-discrimination directives” leads, in the American definition, to unfairly discriminatory rates. This predicament brings Avraham (2017) to contend that lacking clear normative frameworks to judge insurance fairness, the policy makers are working “in a complete legal mess,” which they would contribute to maintain. It might be argued by contrast that normative principles are in fact bound to the society where they emerge (Melé and Sánchez-Runde 2013; Walzer 1995). However, a close look at the decision of the European judge shows in fact that she rejects sex as a parameter for rating because it does not properly reflect the individual risk, thus converging with the American regulatory position.

Yet this common standpoint seems to reduce insurance fairness to an objective, technical/scientific issue. We would like therefore to question in what follows the fairness of actuarial fairness when presented as scientifically reflecting both the risk of the individual insured and the insurer’s technical cost of risk (hence its fairness). By “insurer,” we mean here the insurance practitioner that constructs the fairness of his rates, and not the community of

insured persons that constitute this portfolio. While insurance mechanisms are intrinsically collective and build a cooperative benefit (Baker 2002, 6; Heath 2007, 129; Lehtonen and Liukko 2015, 158), the striving for the adjustment of rates to individual risks in the name of fairness currently blurs this collective viewpoint in favor of an individualistic one. Our main argument is that it derives in fact from a conflation: the actual cost of a group of insureds from the perspective of the insurer (or portfolio manager) is sound from a business profitability viewpoint, but it is arguable whether it represents the “fair price” for the individual insured. We thus hope to help re-center the debate around insurance fairness on its underlying solidarity mechanisms rather than technical and actuarial considerations.

The first section aims at showing in a historical perspective the intertwining of conceptions of fairness with knowledge, in order to point to the alternative to actuarial fairness. The second and main section describes the conflation and shows the intrinsic difference between the insured and the insurer when underwriting an insurance contract. The last part builds on this distinction to discuss the meaning of fairness in insurance prices.

1. Insurance Fairness and Knowledge

There are at least two conceptions of fairness in insurance. As Landes (2015, 519) rightly remarks, one should distinguish between “insurance as the general cooperative arrangement among individuals based on risk pooling and specific arrangements where an insurer acts as an intermediary between policyholders.” The distinction leads to two different definitions of fairness (Ericson, Doyle, and Barry 2003, 6): one focused on the cooperation between the members of the pool, the other on the equity of the contract between the insurer – or the manager of a portfolio of insureds- and the insured. Historically, we want to argue, the two conceptions did not emerge at the same time and reflect

different epistemological and political perspectives.⁶⁴ We would like in this part to describe each approach and the epistemological assumptions that support them, as this will serve our main argument in the next part.

1.1 A Historical Overview of Insurance Fairness

In her seminal study, Daston shows how insurance - together with gambling and annuities-, was long condemned as an “aleatory contract,” a trade between “here-and-present certain goods for uncertain future goods” (Daston 1987, 238; Zelizer 1979). In the seventeenth century, the first generations of mathematical probabilists tried to establish the equity of these contracts on the basis of their mathematical expectation: the fair (equal) price for each participant in a lottery was determined by the totality of possible gains divided by the number of participants (Daston 1987, 240; Martinez, Teira, and Pradier 2016). Fairness was thus understood within the framework of individual equitable contracts.

By the nineteenth century, things had changed. Following Foucault (2004), we argue that a new object of knowledge had taken shape in the form of the population, as another level of reality than the individual one. Insurance mechanisms, now understood as a moral activity rather than gambling,⁶⁵ offered a novel means to cope with uncertainty; while at the individual level accidents remained unpredictable, some regularity appeared at the collective level. The pooling of risks thus allowed for their compensation: accidents, once conceived as the inescapable divine design, became somehow governable. In the coping with adversity, individual prudence -for long the sole remedy-, was replaced by mutualization (Ewald 1986).

⁶⁴ For Baker the two conceptions also reflect different cultural backgrounds (see Baker 2002); for Lehtonen and Liukko, recent developments in marketing that downplay mutuality and solidarity in insurance contribute to the current predominant conception of insurance as a bilateral contract (Lehtonen and Liukko 2015).

⁶⁵ Daston insists that the changed perception towards insurance did not come from new technological capacities but from new social and cultural values (Daston 1987).

This change of perspective, from the individual to the group, was accompanied by a novel vision of equity. The kind of fairness introduced by modern insurance can indeed be grasped at three levels: conceptually first, it stems from an assumption of universal equality in condition in front of the accident. If the latter is not the will of God but a random event, then all are equally exposed. Second, and as a consequence, the moral responsibility for the accident is not looked for in the behavior of the faulty individual but attributed collectively to society: as Baker puts it, insurance “socializes responsibility” (Baker 2002, 2). This later led to the emergence of social rights to the side of liberal natural rights, and the establishment during the twentieth century of welfare states, also coined “insurance societies” (Ewald 1986). With these mechanisms, some of the financial responsibility for the consequences of accidents is covered by insurance.⁶⁶ Insurance, in this description, is understood most generally as the technique of pooling at the heart of the new apprehension of risk, which creates a third level of fairness: the losses, understood as belonging to the collective, are also bore collectively (Lehtonen and Liukko 2011, 35).⁶⁷

The modern, collective concept of fairness did not replace the individual equitable contract though, but somehow co-existed with it; according to Thiery and Schoubroeck (2006), it is mostly upheld today by lawyers who defend an individualistic approach to equality. Actuarial fairness, we would like to add, that posits that “each insured should pay for the risk he brings to the pool,” also belongs to this strand of thought.⁶⁸ In this perspective, insurance is

⁶⁶ Baker (2002) distinguishes between five kinds of responsibility: responsibility as trustworthiness and autonomy are less relevant to our endeavor here. Causal responsibility is involved in the moral responsibility of the accidents. The transfer of burden to society, at least of work accidents, is visible for instance on the no fault systems adopted in work compensation insurance at the end of the nineteenth century (Ewald 1986). Collective responsibility as social accountability and solidarity also result from this transfer.

⁶⁷ Although, as will be discussed below, the distribution of the cost between the members of the group will determine the level of solidarity embedded in the pooling.

⁶⁸ According to Thiery and Schoubreck, actuarial practices follow the aggregate approach. This is true as long as the information is scarce; however, as they also notice, insurers aim to

understood as the transfer of an individual's risk to an insurer, thus downplaying the pooling between insureds: "since the insurer assumes the individual insured's risk of loss, the premium should be fundamentally based upon the expected value of an insured's losses" (Walters 1981, 3). Walters further claims that insurance has actually nothing to do with solidarity, understood here as the subsidizing of some by others:

Some may believe that insurance is an instrument of social policy to compensate victims. This view treats the premiums as merely a means of accumulating funds to pay out losses in ways *possibly fundamentally different from the relative risk that each insured presents to the pool*. But trying to do something noble via the premium collection facilities of insurers *does not make the resultant mechanism insurance*, as cited earlier. Insurance is what it is- the transfer and reduction of risk; it is not a tax to redistribute wealth (Walters 1981, 5, emphasis added).

Since the last decades of the twentieth century, the modern welfare states have entered a period of crisis, due first to financial difficulties (Rosanvallon 1992, 21) that later turned into ideological ones: neoliberal principles questioned the intervention of the state in any social or economic activity, including insurance (Ericson, Barry, and Doyle 2000; Rosanvallon 1995). The individualistic approach took precedence over the collective one, thus giving higher rank to fairness as the equity of a contract and its actuarial principles.

1.2 Epistemological Assumptions behind Actuarial Fairness

Interestingly referring to Rawls, Ewald suggests that at their inception, pooling mechanisms relied on the assumption of a "veil of ignorance" concerning the true nature of the risk of each individual (Ewald 2014, 8); yet while Rawls imagined this veil as a stratagem to erase the known position of each in society (Rawls 2005), the veil that Ewald has in mind was structural. In nineteenth century, the observed regularity of random events at the population level blurred the individual one. Foucault explains how these mechanisms are

decrease subsidizing solidarity as much as possible through increasing classification (Thiery and Schoubroeck 2006, 195–96).

distinct from discipline, precisely because they do not intend to control each individual body, but manage the phenomenon at the aggregate level (Foucault 2004, 63–64).⁶⁹ The newly gained statistical knowledge of the population implied to accept the opacity and unpredictability of the phenomena at the individual level. Hence the fairness of insurance mechanisms relied on a “constitutive ignorance” as concerns the measure of individual risks (Ewald 1986, 368) and allowed for the conception of a new form of solidarity, enacted in the insurance principles (Lehtonen and Liukko 2011).

The individualistic shift towards actuarial fairness has drastically different epistemological assumptions, leading to a series of controversies and paradoxes. By encouraging the actuary to strive for rate adequacy in each and every segment of his portfolio, it indeed assumes that the underlying risks are not equal and that they are knowable. In this perspective, the identification of “persons with risks of a like kind and hazard,” at the heart of the fairly discriminatory rates, becomes the core of the insurance practice. Segmentation of insurance portfolios therefore increased over time. In automobile and property insurance for instance, the classification was enlarged to include parameters such as occupation, age, sex, marital status, level of education, occasional operation of a car, parenthood, and territory (Austin 1983, 536-545), reaching thousands of segments today with the widespread use of generalized linear models (Anderson et al. 2007).

In 1983, Austin reports with irony the following segmentation for youthful drivers:

At the present time the youthful driver categories generally include the following subgroups: single females under twenty-five who are occasional operators of the insured vehicle; single female principal operators under twenty-five; married males under twenty-five; single

⁶⁹ Foucault gives as an example the changing approach to diseases and the type of rationalization implied by vaccination: one knows that some will die from the vaccine, but on overall, in the aggregate, the population is better off than without any vaccination (Foucault 2004, 61–62).

male occasional operators under twenty-five; and single male principal operators under twenty-nine (Austin 1983, 539).

Yet pricing specific groups (provided they are large enough) on the average is already not exactly the same as applying the latter to the entire population. Although it can be viewed as a group approach, it also introduces some differentiation between groups within the population; the assumption of universal equality is lifted, paving the way for more granular segmentation, and its questioning on the side of the insureds. By the end of the 1970s, a series of legal and public controversies emerged around insurance classification (Austin 1983), with parallel debates among actuaries (Lemaire 1984).

Insureds indeed filed complaints for being unfairly affected to a specific (high risk) group. A first family of arguments tackles the classification on the ground that it does not reflect knowledge but social prejudices, masked in mathematical formulas (Austin 1983, 537–50; Glenn 2000); it thus seems to reaffirm the impossibility of knowing individual risks - for long the traditional hypothesis in insurance. The second series of claims considers that any categorization, however statistically justified, is always arbitrary as it never renders the exact characteristics of the individual: here by contrast the individual risk is posed as having a materiality that the classification does not render properly. Actuarial fairness by contrast assumes that some (all?) knowledge on the individual risks is accessible, either to the actuary or at least to the insured. While this ideal situation of total knowledge was purely theoretical to start with, growing technological capacities and available data requested further segmentation, both from scientific and customers' perspectives:

In the past, it used to be very difficult to discover risk factors both in a qualitative and in a quantitative sense (...) But recent developments have changed this situation: with the help of computers it has become possible to make thorough risk analyses, and consequently to arrive at further premium differentiation; the consumer's attitude towards tariffs became more critical. He requested more information and, if he was a good risk, objected to pay the same

premium as the bad ones (De Witt and Van Eeghen 1984, 155, emphasis added).

Yet the parameters that should or could be taken into account constitute a further point of debate; some argue that no parameters should be banned a priori; Walters (1981, 17) thus claims that “one should not legislate against the use of knowledge in a free society.” Others consider that risk parameters that are not under the control of the insured (such as age or gender) should be avoided (Landes 2015; Lehtonen and Liukko 2011; Austin 1983; Thiery and Schoubroeck 2006). Others still would like to introduce socioeconomic parameters that, depending on the line of insurance, do not necessarily reflect risks, in order to foster social fairness (De Witt and Van Eeghen 1984).

Finally, with the attempt to consolidate actuarial fairness principles, the asymmetry of information became a major issue (Akerlof 1970; Stiglitz 1975; Crocker and Snow 2000), with strategies aiming at revealing the insured risk, through “self-selection” for instance (Feldman and Dowd 2000; Finkelstein and Poterba 2004). This discourse assumes that the insured has some knowledge about his risk that the insurer has not. In such a case, if the rates do not follow the principle of actuarial fairness, the insurer faces adverse selection, either because the higher risks will rip a benefit from the process of insurance (Landes 2015, 521), or because in a competitive situation, other insurers might attract the better risks at lower rates (Charpentier and Denuit 2015; Heath 2007; Liukko 2010). In both cases, the non-actuarially fair pricing will result in an underpriced book of business, hence turning problematic for the insurer’s solvency: it might then be argued that actuarial fairness is actually a necessity reflecting the insurer’s needs and viewpoint rather than fairness (Thomas 2007). For Liukko, adverse selection is in fact used to justify derogations to important anti-discrimination principles: “risk classification is deemed such a vital requirement for the sustainable working of private insurance that it is an acceptable reason for discrimination, even to the extent where such

discrimination can be prohibited in other spheres of economic life” (Liukko 2010, 463).

1.3 Two Conceptions of Fairness

Arguments against classification continue today to take two opposite forms, revealing the two competing conceptions of fairness, usually supported by different actors. There are on the one hand those already found in the 1970s controversy, where insureds demand not to be mutualized within a specific group, since their own, individual characteristics show, in their view, a lower incidence. They thus ask for a better adjustment to their own profile. This individualistic approach can be found most recently in the “Test-Achats case” brought to a European court: speaking of life insurance, it was claimed that the use of sex in pricing did not properly reflect the risk as “life expectancy [...] is strongly influenced by economic or social conditions as well as by the habits *of each individual*” (Pinsent Masons 2012).⁷⁰

These arguments seem therefore to strengthen the actuarial fairness standpoint. One should notice that with actuarially adjusted rates the pooling and solidarity do continue to exist, but they are ideally limited to mutuality among insureds of the exact same risk; the constitutive ignorance is reduced to the pure probabilistic, random occurrence within an otherwise perfectly homogeneous group that share a “chance solidarity” (De Witt and Van Eeghen 1984; Lehtonen and Liukko 2011; Thiery and Schoubroeck 2006; Liukko 2010). In extreme cases though, those classified as very high risks might find the burden of insurance too high or face being “redlined;” the actuarially fair premium might indeed lead to the exclusion of some (Lehtonen and Liukko 2011).

An argument in favor of more mutualization and solidarity on the other hand usually comes from the states; the European regulation against the use of

⁷⁰ And actually, when explicitly emphasized by a policyholder, this argument is generally nothing but a way to express the desire of paying not “a fair price”, but a lower price.

genetic information passed in the 1990s for instance aimed at protecting the public at large and avoiding the exclusion of high risk profiles from health insurance (Ewald 2014, 9; Rothstein 2018). In the same strand of thought, the original 2004 European directive against the use of sex aimed at ensuring equal treatment of men and women in goods and services. Its application to insurance, delayed by a derogation, might be taken to foster more solidarity between sexes. Yet the 2011 guidance to insurers for the final implementation of the directive refers to other arguments, that blur the traditionally collective approach of the regulator:

At the moment, a careful young male driver pays more for auto insurance *just because he is a man*. Under the ruling, insurers can no longer use gender as a determining risk factor to justify differences in *individuals' premiums*. But the premiums paid by careful drivers – male and female – will continue to decrease *based on their individual driving behavior* (European Commission 2011, emphasis added).

The text asks for more granular data than gender categories, thus pointing to individual behavior rather than an aggregate approach. Such an argument in 2011, when telematics makes individual driving habits available, cannot be neutral. The allusion seems to shift the position of the state from an approach that was traditionally group based, toward an individualistic one, one that was the apantage of individuals tackling classification. It seems thus to support a pricing that would be based on individual behavior, further weakening the conception of fairness as broad solidarity.

In the neoliberal context, this trend is not surprising; Karsenti and Lemieux show that each ideology is accompanied by its own knowledge; hence the welfare state was relying on sociology and the vision of society as a whole, whereas liberalism and neoliberalism make use economics and methodological individualism (Karsenti and Lemieux 2017).⁷¹ Yet our main argument here is that there is a basic flaw when using the expected value of loss as a principle of

⁷¹ Rawls might be an exception here, as he tried to legitimate the group approach with methodological individualism; the veil of ignorance allows him to re-construct the Kantian “enlarged mentality”, based on the individual alone (Rawls 2005)

fair pricing, as will be exposed in the next part. The point is all the more important today as the group approach to fairness, as we have tried to show above, is being downplayed by most actors of the industry.

2. Plurality, Randomness and Actuarial Fairness

In the two approaches to fairness evoked in the previous part, the available knowledge seems to play a crucial part; once some information becomes available, the assumption of equality that accompanies the group approach is lifted and replaced with risk differentiation, thus shifting fairness to an actuarial, individualistic definition. Yet the knowledge at stake depends, we would like to argue here, on two distinct uncertain situations that have further consequences to the understanding of insurance fairness.

2.1 Plurality vs. Randomness: a Conceptual Framework

At the foundation of insurance lies the existence of uncertain events. An agent is in an uncertain situation when he has to consider a future phenomenon with multiple possible results. The possible results themselves might be known to the agent, but the actual realization that will occur remains unknown. Notwithstanding the fact that both the insurer and the insured face the same uncertain event, the insured's and the insurer's situations are not identical, as the policyholder is exposed to one single realization, whereas the insurer signs a large number of contracts and observes – in most types of insurance- a significant amount of events. Taken as a whole, thanks to the law of large numbers, the insurer can indeed forecast his final results on the aggregation of contracts. If the law of large numbers applies, the insurer's portfolio will have an almost deterministic result. The uncertainty of the result is thus much smaller for the insurer, for whom the future is more or less predictable, than for the insured that remains in an radically unknown position.

There is therefore a difference in the nature of the situation, depending on whether the underlying phenomenon repeats itself for the observer: in one case –the single event-, the situation remains random, whereas in the other –the plurality of occurrences- it is almost deterministic. In the lines of insurance

where the law of large numbers applies, the insurer is in a plural situation: he will face a high number of occurrences, at least high enough so that the deviation on the aggregated result becomes negligible from his viewpoint. The plurality situation and the type of calculation it allows appear only at the aggregate level evoked in the previous part. The regularities that led to the conceptualization of the population are indeed a consequence of a situation of plurality.⁷²

The insured by contrast chooses to buy or not to buy the contract without knowing what will occur: he will make this kind of decision a few times in his life, and as such he is and remains in a random situation. Following Knight's terminology, the insurer faces a risk that is measurable, whereas the insured remains in pure uncertainty (Knight 1985).⁷³ In this terminology, the insurance contract is not a transfer of risk, but a transfer of uncertainty that, once transferred to the insurer, becomes a measurable risk. Hence when speaking of equitable contracts, one should take into account this basic difference between the insured and the insurer, and the fact that they do not face the same uncertainty.

Two further points are to be highlighted here; first, the type of the situation is contingent upon the observer: the same event belongs to a random and/or plural situation depending on the perspective. Second, the distinction between a random situation and a plural situation derives solely from the number of future realizations of the phenomenon; the more or less exact knowledge of the event's mathematical probability does not matter. This point is most of the

⁷² Yet the collective level does not by itself grant a situation of plurality, as the latter really depends on the number of observed occurrences, as will be shown below on the example of catastrophes.

⁷³ Knight (1921, pp. 224-225) distinguishes three types of situations: *a priori* probabilities, which are purely mathematical as in a dice throw, for example; *statistical* probabilities, or the empirical observed frequency within a class of events considered to be homogeneous; and estimates, for which there is "*no valid basis of any kind* for classifying instances" [his italics]. In accordance with our analysis above, we consider that the distinction between these last two categories does not necessarily depend on the intrinsic nature of the phenomenon but also on the observer's position relative to the phenomenon.

time forgotten in disaster planning (i.e. random situations), where risk mitigation is attempted through the projection of a known past on an irreducibly unknown future (Cevolini 2012, 365). In cases where the insurer is himself in a random situation, the use of expected values shows only how the “probabilistic approach to risk” has become prevalent in the last century, despite its being controversial (McGoun 1995).

The distinction between plurality and randomness thus allows pinpointing a fallacy that frequently occurs in the insurance pricing literature. Actuarially fair pricing is indeed based on the expected value of the uncertain event at stake, taken to be “the risk” transferred by the insured to the insurer, as if the insured had the possibility to make the same kind of calculations as the insurer: “since the insurer assumes the individual insured’s risk of loss, the premium should be fundamentally based upon the expected value of an insured’s losses. *The expected loss for an insured is the average or probable number of losses (or claims) times the average cost of those claims*” (Walters 1981, 3, emphasis added).

Yet the expected value of the losses does not stem from the insured’s viewpoint, but solely from the insurer’s: it cannot be explained intuitively without relying on the notion of average, that is to say on an estimator that could be observed and measured were the game to be played several times. Even in a thought experiment, one cannot interpret the concept of expected value otherwise than as embodied in an average, i.e. by referring not to a single event, but to a whole set of events. In other words, the expected value makes sense in a plural situation, but the very definition of a random situation contradicts the reasoning that consists in imagining a repetition of the event.

Furthermore, when the number of future realizations of a phenomenon is low for a given observer, as is the case in a random situation, the total number of observations (both past and future) for this observer generally remains low. In such a case, one faces the “reference class issue,” i.e. the difficulty or the impossibility of constituting a sample of realizations of the same uncertain

phenomenon, necessary to the calculation of an expected value (McGoun, 1995; Eagle 2004). As von Mises (1949) pointed out, a unique event is never a member of a class, and therefore, in a random situation, an objective measure of the individual expected value is impossible to obtain.

Hence while the insurer's plural situation allows for the calculation of an expected value that properly reflects his position and measures his risk, the policyholder remains by contrast in a random situation that intrinsically invalidates such a calculation. The conflation between the insurer's plurality position and the insured's randomness is widespread in current literature on risk; it is a secondary result of the adoption of a probabilistic approach to risk, followed by new regulations that make it normative (McGoun 1995; François and Frezal 2018; Frezal 2018).

This distinction between plurality and randomness allows us to introduce further criticism of actuarial fairness on two main points: its use of expected value as a measure for the individual, and the faith that further segmentation leads to the convergence towards the "true" probability of the individual event.

2.2 On expected value as the fairness reference framework

The assumption that expected values are or should be used by rational individuals for taking decisions in uncertain situations is grounded in game theory as developed by von Neumann and Morgenstern (1944), further reviving a formulation attributed to Daniel Bernoulli (see also Friedman and Savage 1948). While it has become a frame of reference in many domains, such as insurance fairness, it has been tackled almost since inception. One of its early criticism was voiced by Allais, who noticed that the expected utility intuitively relies on the law of large numbers. However:

It is obvious that, if it is a question of a single dice throw that will not recur, the justification founded on the law of large numbers does not apply at all. Yet, and precisely, most cases where we have to make random decisions are isolated cases. There, in our sense, the law of large numbers has not any utility. (Allais 1953, 543, original emphasis, our translation).

Allais further notices that the maximization of utility principle remains unclear on whether it should be understood as a normative framework – how rational agents ought to behave, or whether it describes actual behaviors of supposedly rational individuals. However, none of them holds (Allais 1953).

Further criticisms of rational decision theory led to the emergence of behavioral economics as a new way to describe actual behaviors (A. Kahneman and N. Tversky 1979). Applied to insurance, it confirms that insureds do not behave according to the utility maximization principle (Baker and Siegelman 2013). Decision theorists have since come to the conclusion that no model built on expected utility can successfully combine the necessary answers to the different observations in one global theory.⁷⁴ For Mathis and Steffen, the main difference between rational choice theory and behavioral economics is precisely the fact that although the former might inaccurately describe specific individual actions, it gives very good results when looked at on the overall; its aim is to describe the population and not the individual (Mathis and Steffen 2015). This confirms that the expected value is not an indication for the individual, but only makes sense when looked at on the aggregate; in our case this means that it describes properly the viewpoint of the insurer alone.

The paradox is thus as follows: the group approach that fostered insurance fairness as welfare and solidarity actually rendered possible the emergence of “risk” as measured on the expected value; however, the individualistic turn that accompanies neoliberalism attempts to use this indicator -that holds only on the aggregate-, for a conception of fairness now focused on the individual. Yet the expected value has no specific legitimacy from the viewpoint of the policyholder. Hence, from the viewpoint of a male driver, it is possible to claim that gender segmentation, although accurately reflecting the costs for the insurer, is discriminatory. From his viewpoint, such a postulated “fair price”

⁷⁴ See for instance Camerer (1989) or Barberis (2013). For concrete and recent examples of observations disagreeing with standard expectation based theories, see for example Barreda-Tarrazona et al., 2015.

does not correspond to a criterion of justice. However, the expectation that some other premium might fairly reflect “his exact risk” has no basis either, as will be shown below.

2.3 On expected value as an objective measure of the individual “true risk”

Considering that the fair price of insurance is embodied in the individual’s pure premium further relies on another implicit assumption; that the exact individual probability of claim exists and can be asymptotically reached (De Witt and Van Eeghen 1984). However, the segmentation refinement fostered by actuarial fairness does not warrant that the individual pure premium will converge towards a “true” value, to the contrary. The probability, for a given person, to suffer from a specific hazard can be calculated based on a series of parameters, all depending on the level of knowledge and data available in a given time and place. Historically, the probability over the global population was first made available; later, demographic parameters such as age and gender were introduced. As a result, the probability became lower for some, and higher for others. Today, information about existing medical conditions is available and might be considered as a parameter for pricing. Again, the additional parameters will further discriminate within the existing groups, thus the probability for a specific individual might become higher or lower. The emergence of new behavioral data provided by connected devices will further create sub-groups with new probability levels, some higher and some lower, and so forth with any new piece of collected data.

Actuarial fairness thus conceptualized means that the rates contain less subsidies between groups and that premiums are converging towards a supposedly “optimal” level, with no subsidies at all (De Witt and Van Eeghen 1984). The transposition to the individual, at the heart of the conflation between plurality and randomness, means that the additional parameters would marginally change the level of his premium. In practice, each step in the segmentation changes his risk estimator, but the refined granularity on the overall does not imply that his specific estimator is getting closer to any

specific value, nor that the amplitude of the adjustments decreases. In other words, increasing segmentation does not make the individual's premium converge towards any "real" cost of risk.

3. Discussion: Fairness in Insurance

Insurance is a social construct; as Baker puts it, it defines the limits of individual and collective responsibility. By setting the level of solidarity that is collectively accepted, from which fairness derives, it thus shapes the society we live in (Baker 2002). It is sometimes argued that the two kinds of fairness evoked in the first part reflect two levels of solidarity. The group approach to fairness would imply a "subsidizing solidarity," where subsidies between different risks are accepted. Actuarial fairness and its individualistic approach by contrast strives for pure "chance solidarity," where solidarity continues to exist between people with the exact same risk characteristics, that share with one another the sole randomness of occurrence (Lehtonen and Liukko 2011). Yet we have tried to show here that this definition of "true" homogeneity within a group is at best an illusion. The definition of what is shared between risks as "pure chance" is contingent upon the knowledge available in a given time and place, and simply reflects what is not expressed by the classification. Negotiating the limits of a classification is therefore not an objective and scientific exercise; it is an arbitration between economic efficiency and social justice, that sets the balance between exclusion from insurance faced by higher risks on the one hand and the insurer's exposure to moral hazard and adverse selection on the other hand (Avraham 2017).⁷⁵

For Avraham (2017), the leading principles when deciding on the legitimacy of parameters should be both their social fairness (considering the soundness of the discrimination at stake) and their impact on justice, i.e. the social and economic consequences of their prohibition or adoption: prohibiting

⁷⁵ See Dionne and Rothschild (2014) for a thorough analysis of the moral hazard issue and (Baker 1994) for a historical approach.

classification based on genetic tests for instance might lead, via adverse selection, to the exclusion of the good risks from insurance and higher rates for the insureds (Avraham 2017). But this classification can also have negative social effects as it might prevent people from taking genetic tests lest their insurance be declined, thus denying them a proper treatment of their condition (Rothstein 2018).

Liukko distinguishes between, on the one hand market and industry justifications - both in favor of some classification for the sake of competition and efficiency respectively - and civic justification on the other hand that puts to the fore some vision of the common good. The level of solidarity in insurance results from a compromise between those distinct and sometimes contradictory goals (Liukko 2010). In any case, insurance fairness should be understood as an intrinsic part of social justice. Its determination should not be blurred under spurious scientific indicators.

Conclusion

Insurance fairness can be understood either as the solidarity among a group of insureds pooling their uncertain results, or as a fair contract between an insured and an insurer. In the course of the twentieth century, the growing influence of neoliberal theories has downscaled the first approach, that accompanied the rise of welfare states, in favor of the individualistic approach. This shift makes of “actuarial fairness” the central conception of fairness in insurance. This conception assumes that the pure premium represents a common indicator of the risk for both the insurer and the insured, thus being a sound and fair base for pricing. However, we have tried to show that the economic and financial tools for apprehending risks on the basis of statistical indicators are founded on the conflation of two phenomena: plurality, for which these statistical concepts have emerged, and randomness, onto which these concepts have been wrongly transposed.

The pertinence of certain tools is indeed contingent on the observer's position in respect to the phenomenon he seeks to understand. The transposition onto random situations of tools conceived for understanding plurality is exceedingly easy from the standpoint of mathematical formalism, but it is based on the identification of two decision-makers whose positions are radically different. If, for observers facing plurality (e.g. insurers), expected values provide meaningful insights into decision-making, for observers in a random situation (e.g. insureds), this transposition is invalid because it does not properly describe his situation.

As exemplified in the recent debates over legitimate parameters for classification, common visions over fair-pricing regulations are focused on expected value, and are therefore subject to these failings. We have tried to show in the above that in considering the fairness of an insurance scheme, the fairness of parameters cannot be measured on their being adjusted "to the individual's risk." As a consequence, we suggest to re-center the discussion around issues of social fairness, without sacrificing these on the altar of a fallacious technico-ethical customers' equity.

We have tried to suggest that these discussions should remain in the field of the political choices gathering economic efficiency (adverse selection vs. robustness) and societal priorities (freedom of insurers vs. solidarity): it shall not mobilize the frequent but fallacious argument of the actuarial fairness of paying one's "own risk."

References

- Akerlof, G. A. (1970). "The Market for Lemons : Qualitative Uncertainty and the Market Mechanism." *Quarterly Journal of Economics* 84 (3): 488–500.
- Allais, M. (1953). Le comportement de l'homme rationnel devant le risque : critique des postulats et axiomes de l'école américaine. *Econometrica*, 21 (4) : 503-546.

- Anderson, D., Feldblum S., Modlin C., Schirmacher D., Schirmacher E., and Thandi N. (2007). "Actuarial Insurance Generalized Linear Model Practical Guide." Willis Towers Watson. 2004 2007. <https://www.towerswatson.com/en/Insights/IC-Types/Technical-Regulatory/2010/A-Practitioners-Guide-to-Generalized-Linear-Models>.
- Austin, R. (1983). "The Insurance Classification Controversy." *University of Pennsylvania Law Review* 131 (3): 517. <https://doi.org/10.2307/3311844>.
- Avraham, R. (2017). "Discrimination and Insurance." SSRN Scholarly Paper ID 3089946. Rochester, NY: Social Science Research Network. <https://papers.ssrn.com/abstract=3089946>.
- Baker, T. (1994). "On the Genealogy of Moral Hazard." *Texas Law Review*, 59.
- Baker, T. (2002). "Risk, Insurance, and (the Social Construction of) Responsibility." *University of Connecticut School of Law Articles and Working Papers*, January. http://lsr.nellco.org/uconn_wps/8.
- Baker, T. and Siegelman, P. (2013). Behavioral Economics and Insurance Law: The Importance of Equilibrium Analysis. Published in: Teichman T. & Zamir E. (eds.), *Oxford Handbook of Behavioral Economics and the Law* (2014). Available at SSRN: <https://ssrn.com/abstract=2386158>
- Barberis, N. (2013). The Psychology of Tail Events: Progress and Challenges. *American Economic Review*, 103 (3) May 2013: 611-616.
- Barreda-Tarrazona, I., Jaramillo-Gutierrez, A., Navarro-Martinez, D. and Sabater-Grande, G. (2014). The Role of Foregone Opportunities in Decision-Making Under Risk. *Journal of Risk and Uncertainty*, 50 (2): 167-188.
- Camerer, C.F. (1989). An Experimental Test of Several Generalized Utility Theories. *Journal of Risk and Uncertainty*, 2 (1): 61-104.
- Cevolini, A. (2012). Coping with the Unknown – Time and Technology in Formal Organizations Searching for Safety. *Distinktion: Scandinavian Journal of Social Theory*, Vol. 13, No. 3, December 2012: 355–371.

- Charpentier, A. and M. Denuit. (2015). “Segmentation et Mutualisation, Les Deux Faces d’une Môme Pièce.” *Risques*.
https://www.researchgate.net/publication/282723284_Segmentation_et_Mutualisation_les_deux_faces_d'une_meme_piece.
- Consumer Federation of America (CFA). (2013). Letter sent 29 August 2013 by R. Hunter, Director of Insurance, Casualty Actuarial Society.
- Court of Justice of the European Union (CJEU). (2011). 1 March 2011 Judgment in case C-236/09, curia.europa.eu.
- Crocker, K. J., and A. Snow. (2000). “The Theory of Risk Classification.” In *Handbook of Insurance*, 245–76. Huebner International Series on Risk, Insurance, and Economic Security. Springer, Dordrecht.
https://doi.org/10.1007/978-94-010-0642-2_8.
- Daston, L. (1987). “The Domestication of Risk: Mathematical Probability and Insurance, 1650-1830.” In *The Probabilistic Revolution, Vol 1: Ideas in History*, edited by Lorenz Krüger, Lorraine Daston, and Michael Heidelberger, 237–61. Cambridge: The MIT Press.
- Daston, L., (1988). *Classical Probability in the Enlightenment*, Princeton University Press, New Jersey.
- Daston, L., (1989). L’interprétation classique du calcul des probabilités. *Annales, Economies, Sociétés, Civilisations*. 44 (3): 715-731.
- De Witt, G.W., and J. Van Eeghen. (1984). “Rate Making and Society’s Sense of Fairness.” *ASTIN Bulletin*, no. 14:2: 151–64.
- Dionne, G. and Rothschild, C. (2014). Economic Effects of Risk Classification Bans. *Geneva Risk and Insurance Review*, 39 (2): 184-221.
- Doherty, N. (2013). Risk and the Endogenous Economist: Some Comparisons of the Treatment of Risk in Physics and Economics. *Geneva Risk and Insurance Review*, 38 (1): 1-22.
- Eagle, A. (2004). Twenty-One Arguments Against Propensity Analyses of Probability. *Erkenntnis*, 60: 371-416.

- Ericson, R.V., D. Barry, and A. Doyle. (2000). "The Moral Hazards of Neo-Liberalism: Lessons from the Private Insurance Industry." *Economy and Society* 29 (4): 532–58.
- Ericson, R.V., A. Doyle, and D. Barry. (2003). *Insurance as Governance*. University of Toronto Press.
- European Council. (1997). Convention for the Protection of Human Rights and Dignity of the Human Being with Regard to the Application of Biology and Medicine: Convention on Human Rights and Biomedicine (Oviedo), arts 11 and 12, <https://www.coe.int/conventions>
- European Council. (2004). Directive 2004/113/EC Implementing the Principle of Equal Treatment between Men and Women in the Access to and Supply of Goods and Services. *Official Journal of the European Union*, L373 1, 21.12.04.
- European Commission. (2011). Presse Release - European Commission gives guidance to Europe's insurance industry to ensure non-discrimination between women and men in insurance premiums. Available online at: http://europa.eu/rapid/press-release_IP-11-1581_en.htm.
- Ewald, F. (1986). *L'Etat providence*. Paris : Grasset.
- Ewald, F. (2014). "Assurance-Prévention-Prédiction. Big data et société – Institut Montparnasse." <http://www.institut-montparnasse.eu/nos-publications/les-inegalites-territoriales-et-sociales-2-12/>.
- Feldman, R., and B. Dowd. (2000). "Risk Segmentation: Goal or Problem?" *Journal of Health Economics* 19 (4): 499–512. [https://doi.org/10.1016/S0167-6296\(00\)00051-5](https://doi.org/10.1016/S0167-6296(00)00051-5).
- Feller W. (1968). *An Introduction to Probability Theory and its Applications*, Vol. 1, 3rd ed., John Wiley & Sons Inc., New York.
- Finkelstein, A., and J. Poterba. (2004). "Adverse Selection in Insurance Markets: Policyholder Evidence from the U.K. Annuity Market." *Journal of Political Economy* 112 (1): 183–208.

- Foucault, M. (2004). *Sécurité, Territoire, Population - Cours au Collège de France. 1977-1978*. Paris: Gallimard/Seuil.
- François, P., and S. Frezal. (2018). "Instituer l'incohérence. L'amalgame Aléa et Hétérogénéité Dans Le Secteur Assuranciel." *Sociologie Du Travail* 60 (1): 1-23.
- Frezal, D. (2018). *Quand les statistiques minent la finance et la société: Risque, responsabilité et décision*. Paris : L'Harmattan.
- Friedman, M. and Savage, L. (1948). Utility Analysis of Choices Involving Risk. *Journal of Political Economy* 56 (4): 279–304.
- Gilboa, I. (2009). Is It Always Rational to Satisfy Savage's Axioms?. *Economics and Philosophy*, 25, Special Issue 03: 285–296.
- Glenn, B. J. (2000). "The Shifting Rhetoric of Insurance Denial." *Law & Society Review* 34 (3): 779. <https://doi.org/10.2307/3115143>.
- Heath, J. (2007). "Reasonable Restrictions on Underwriting." In *Insurance Ethics for a More Ethical World*, 7:127–59. Research in Ethical Issues in Organizations 7. Emerald Group Publishing Limited. [https://doi.org/10.1016/S1529-2096\(06\)07007-6](https://doi.org/10.1016/S1529-2096(06)07007-6).
- Jones, D. (2013). Re: Statement of principles Regarding Property and Casualty Insurance Ratemaking. *Letter to the Chief Actuary of Louisiana Department of Insurance*, 2013, May 21.
- Kahneman, D. and Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, 47 (2): 263-292.
- Karsenti, B. and C. Lemieux. (2017). *Socialisme et sociologie*. Paris: Editions de l'Ecole des Hautes Etudes en Sciences Sociales.
- Knight F. 1921. *Risk, Uncertainty and Profit*. Boston and New York, Houghton Mifflin.
- Knight, F. H. (1985). *Risk, Uncertainty and Profit*. Chicago: University of Chicago press.
- Landes, X. (2015). "How Fair Is Actuarial Fairness?" *Journal of Business Ethics* 128 (3): 519–533.

- Lehtonen, T.-K.. and J. Liukko. (2011). “The Forms and Limits of Insurance Solidarity.” *Journal of Business Ethics* 103 (1): 33–44.
- Lehtonen, T.-K.. and J. Liukko. (2015). “Producing Solidarity, Inequality and Exclusion Through Insurance.” *Res Publica* 21 (2): 155–69. <https://doi.org/10.1007/s11158-015-9270-5>.
- Lemaire, J. (1984). “The Influence of Expense Loadings on the Fairness of a Tariff.” *ASTIN Bulletin: The Journal of the IAA* 14 (2): 165–71.
- Liukko, J. (2010). “Genetic Discrimination, Insurance, and Solidarity: An Analysis of the Argumentation for Fair Risk Classification.” *New Genetics and Society* 29 (4): 457–75. <https://doi.org/10.1080/14636778.2010.528197>.
- Martinez, A. J. H., D. Teira, and P-C. Pradier. (2016). “What Was Fair in Actuarial Fairness?” *Documents de Travail Du Centre d’Economie de La Sorbonne*. 2016.73 - ISSN : 1955-611X. 2016. <halshs-01400213>.
- Mathis, K., and A. D. Steffen. (2015). “From Rational Choice to Behavioural Economics.” In *European Perspectives on Behavioural Law and Economics*, edited by K. Mathis, 31–48. Economic Analysis of Law in European Legal Scholarship. Cham: Springer International Publishing.
- Melé, D. and C. Sánchez-Runde. (2013). “Cultural Diversity and Universal Ethics in a Global World.” *Journal of Business Ethics* 116 (4): 681–87. <https://doi.org/10.1007/s10551-013-1814-z>.
- Miller, M. J. (2009). “Disparate Impact and Unfairly Discriminatory Insurance Rates.” *Casualty Actuarial Society E-Forum*, no. Winter 2009. <https://www.casact.org/pubs/forum/09wforum/>.
- McGoun E. (1995). The History of Risk ‘Measurement’. *Critical Perspectives on Accounting*, 6: 511-532.
- Pinsent Masons. (2012). “Test-Achats.” 2012. <https://www.out-law.com/page-11804>.
- Rawls, J. 2005. *A Theory of Justice*. Cambridge, Massachusetts ; London, England: Harvard University Press.

- Reichenbach, H. (1940). On the Justification of Induction. *The Journal of Philosophy*, 37 (4): 97-103.
- Rosanvallon, P. (1992). *La Crise de l'Etat-providence*. Nouv. éd. Paris: Seuil.
- Rosanvallon, P. (1995). *La Nouvelle question sociale. Repenser l'Etat-providence*. Paris: Points.
- Rothstein, M. A. (2018). "Time to End the Use of Genetic Test Results in Life Insurance Underwriting." *Journal of Law, Medicine and Ethics* 46 (3). <https://papers.ssrn.com/abstract=3152572>.
- Savage, L. (1954). *The Foundations of Statistics*, New York, Wiley.
- Siegel, E. (2016). *Predictive Analytics: The Power to Predict Who Will Click, Buy, Lie, or Die*, Hoboken, Wiley.
- Stiglitz, J. E. (1975). "The Theory of 'Screening,' Education, and the Distribution of Income." *The American Economic Review* 65 (3): 283–300.
- Thiery, Y., and C. Van Schoubroeck. (2006). "Fairness and Equality in Insurance Classification." *The Geneva Papers on Risk and Insurance - Issues and Practice* 31 (2): 190–211. <https://doi.org/10.1057/palgrave.gpp.2510078>.
- Thomas, R G. (2007). "Some Novel Perspectives on Risk Classification." *The Geneva Papers on Risk and Insurance - Issues and Practice* 32 (1): 105–32. <https://doi.org/10.1057/palgrave.gpp.2510118>.
- Turow, J., L. McGuiguan, and E. R. Maris. (2015). "Making Data Mining a Natural Part of Life: Physical Retailing, Customer Surveillance and the 21st Century Social Imaginary." *European Journal of Cultural Studies* Vol. 18 (4–5): 464– 478.
- Von Mises, L. (1949). *Human Action, I, VI, 5*, New Haven, Yale University Press.
- Von Neumann, J. and Morgenstern, O. (1944). *Theory of Games and Economic Behaviour*, Princeton, Princeton University Press.
- Walters, M. A. (1981). "Risk Classification Standards." *Proceedings of the Casualty Actuarial Society* 68: 1–23.

Walzer, M. (1995). "Thick and Thin: Moral Argument at Home and Abroad."

Philosophy 70 (273): 472–475.

Zelizer, V. (1979). *Morals and Markets: The Development of Life Insurance in the United States*. New York: Columbia university press.

Partie III : Retour sur les facteurs d'institutionnalisation de
l'approche statistique promue par Solvabilité 2

Chapitre 5 : La réussite d'un amalgame problématique

Instituer l'incobérence. Aléa et hétérogénéité au sein du secteur assurantiel

Introduction

Dans les schémas d'explication avancés par les sciences sociales pour rendre compte de la crise de 2008, certains outils financiers — les produits dérivés, les classements des agences de notation, les outils de prévention des risques — occupent une position déterminante (Lounsbury et Hirsch, 2010). Ces outils, pour la plupart, ont une histoire fort longue, et ce n'est pas à leur seule existence qu'est imputée la plus grave crise financière de l'après Seconde Guerre mondiale. Certains de ces travaux insistent cependant sur la puissance démultipliée que trouvent ces outils dans l'équipement mathématique qui est le leur depuis les années 1970 : selon Akos Rona-Tas (2010) par exemple, ce n'est pas tant l'existence des agences de notation qui explique leur rôle déterminant dans la bulle des *subprimes*, que la confiance aveugle que les analystes pouvaient avoir dans les modèles mathématiques. Bien avant le déclenchement de la crise de 2008, le rôle des modèles mathématiques dans le développement des activités financières a été abondamment étudié, en particulier par des travaux qui mobilisent l'hypothèse d'une *performativité* de ces modèles (MacKenzie, 2006). Selon cette hypothèse, en s'incorporant aux outils cognitifs (les modèles mathématiques) et techniques (les dispositifs informatiques notamment) qui équipent les décisions que prennent les acteurs des marchés financiers, la théorie économique transforme (« performe ») la réalité qu'elle prétend décrire. Comme le soulignent notamment Donald MacKenzie et Taylor Spears (2014), les travaux qui s'inspirent de cette hypothèse s'intéressent davantage aux conséquences de leur mise en œuvre qu'aux processus de leur mise au point. Cette appréciation sévère peut certes être en partie nuancée en tenant compte, par exemple, de l'étude que D. MacKenzie (2003) consacre à l'histoire de la formule de Black-Scholes (voir également Breslau et Yonay, 1999). Il reste que l'essentiel de l'effort théorique des *Social studies of finance* porte sur les effets des modèles, davantage que sur leur mise au point.

L'ambition de cet article est de rendre compte de la genèse de formes de raisonnements que l'on retrouve au cœur des outils de gestion des risques financiers. Le risque que gèrent les institutions financières renvoie à deux classes très différentes de phénomènes. D'un côté, les institutions financières gagnent de l'argent en agrégeant les risques *d'autrui*, par exemple lorsqu'un assureur mutualise les risques de ses assurés et utilise des statistiques pour déterminer le montant de la cotisation qui optimisera sa rentabilité. En mutualisant, l'organisme financier transforme l'*aléa* auquel chacun de ses clients est confronté en une *hétérogénéité* que la loi des grands nombres lui rend déterministe. D'un autre côté, les institutions financières sont confrontées à *leur* propre risque : celui d'un retournement brutal de conjoncture pour une banque par exemple, ou d'une catastrophe sanitaire ou climatique pour une compagnie d'assurance. Ce risque n'est pas déterministe, la loi des grands nombres n'y joue pas et les statistiques ne permettent pas de le prévoir. Nous avons montré ailleurs que les modèles de risque mobilisés par l'industrie financière appréhendent avec le même formalisme mathématique ces deux situations de nature profondément différente, et en quoi cette assimilation peut avoir des conséquences dramatiques (Frezal, 2015a, 2015b). Nous souhaitons ici comprendre pourquoi les acteurs du monde financier appuient leurs raisonnements et leurs décisions en traitant avec les mêmes outils, c'est-à-dire à partir d'indicateurs statistiques, ces deux situations.

Pour rendre compte de cette assimilation des situations d'aléa et d'hétérogénéité, nous faisons l'hypothèse qu'il faut être attentif à leurs propriétés épistémologiques. Les modèles et les hypothèses implicites qui les fondent jouent en effet, dans la sphère financière, le rôle d'une *institution* : ils constituent des cadres *a priori*, normatifs et partagés (voir François, 2011). Ils sont *partagés*, car l'ensemble des modèles mobilisés en mathématiques financières par les acteurs qui s'y réfèrent mettent en jeu ces assimilations. Ils sont *a priori*, car ils ne sont pas discutés : en dépit des critiques qui pourraient leur être adressées, personne ne les remet en cause — comme n'est pas remise

en cause la croyance selon laquelle la danse du sorcier provoquera la pluie, même si après que le sorcier a dansé le soleil continue de briller, pour reprendre l'exemple célèbre d'Henri Hubert et Marcel Mauss (1985). Ces principes, enfin, sont *normatifs*, au sens où les mesures qui en découlent orientent les décisions à prendre : ils précisent s'il faut, par exemple, investir ici plutôt que là. La question qui nous importe n'est donc pas seulement, comme le font D. MacKenzie et T. Spears (2014), de retracer l'histoire de certaines formes de raisonnement — en l'occurrence, l'assimilation entre aléa et hétérogénéité. Elle est aussi de comprendre comment ces formes de raisonnement en sont venues à acquérir ce caractère *institutionnel*.

Pour soulever cette question, nous proposons de nous appuyer sur le cadre conceptuel présenté dans François (2011) — pour une mise en œuvre de ces propositions, voir notamment Dubois et François (2013), Chauvin (2011) ou Ollivier (2011). En nous inspirant de la seconde philosophie de Ludwig Wittgenstein et de la compréhension qu'il y avance de ce que c'est que suivre une règle, nous suggérons qu'aucune proposition, aucune catégorie de pensée, aucun enchaînement causal ne peut par nature se prévaloir des propriétés d'une institution : ce n'est pas en vertu de propriétés intrinsèques qu'une proposition ou une catégorie de pensée en viennent à s'affirmer comme des institutions — ce n'est pas en elles qu'il faut aller chercher ce qui fonde leur caractère institutionnel, mais dans les pratiques qui les mobilisent. Une institution, autrement dit, n'existe que pour autant — dès lors, et aussi longtemps — que des acteurs en usent comme telle. Retracer la genèse d'une institution impose dès lors de reconstituer la généalogie de ses usages : pour comprendre comment naît une institution, il faut saisir comment des acteurs en viennent à faire jouer un rôle normatif et *a priori* à des propositions qui longtemps n'ont pu circuler qu'à titre d'hypothèse.

Ce cadre analytique doit nous permettre de faire retour sur les deux principales tendances que la littérature néo-institutionnaliste a placées en son cœur pour analyser le changement institutionnel (Hwang et Powell, 2005). La première

insiste sur le rôle des *entrepreneurs institutionnels* qui voient dans la mise en place de nouvelles institutions une manière de défendre leurs intérêts, et dans les ressources dont ils disposent les moyens de s'assurer du succès de leur entreprise (DiMaggio, 1988). L'analyse doit alors se concentrer sur les moments instituants où émergent des coalitions d'acteurs visant à promouvoir l'adoption de telle ou telle norme, ou au contraire à en interdire l'avènement (Rao *et al.*, 2003). Comprendre les ressorts et les modalités de ces moments de mobilisation est une étape nécessaire pour rendre compte des dynamiques d'institutionnalisation, mais l'attention à ce que Walter Powell et Jeannette Colyvas (2008) appellent les « microfondations des institutions » a mis en lumière l'importance de processus plus lents, plus souterrains, mais non moins essentiels dans l'avènement de nouvelles institutions : c'est l'intuition que défendent les travaux qui portent sur le *travail institutionnel* (Lawrence et Suddaby, 2006). Les phases de mobilisation et d'explicitation sont en effet souvent préparées, en leur amont, par une longue période d'appropriation par certains acteurs des principes dont ils prendront plus tard la défense : la phase de mobilisation n'intervient souvent qu'après qu'une partie au moins des acteurs s'est approprié les principes dont ils se feront les hérauts (Garud *et al.*, 2002). Une institution, par ailleurs, n'existe pas seulement d'être proclamée telle : il faut que les acteurs lui fassent jouer ce rôle. Il importe alors de saisir les mécanismes qui peuvent porter les acteurs à reconduire certaines pratiques et qui, en les stabilisant, font ainsi jouer à ces règles, à ces principes ou à ces catégories, le rôle d'une institution (Oliver, 1992).

Pour rendre compte du processus d'institutionnalisation d'un raisonnement, il faut donc repérer, d'un côté, comment certains acteurs parfois se mobilisent, et de l'autre comment ils s'approprient, au long cours, des formes de raisonnement qui orientent leurs pratiques et qu'ils ne remettent pas en cause. Mais de quels acteurs parle-t-on, et sur quelles scènes sociales interviennent-ils ? Retracer la trajectoire d'institutionnalisation de l'assimilation des situations d'aléa et d'hétérogénéité dans les modèles de mathématiques financières

impose *a priori* d'ouvrir considérablement l'espace social et chronologique de l'investigation, puisque cette trajectoire se joue, de manière grossièrement séquentielle, sur deux scènes distinctes : celle de ses usages théoriques, dans l'espace social où s'écrit et se développe la théorie des probabilités et les usages qui en sont faits en mathématiques financières et en économie, et celle de ses usages pratiques, dans les organisations et les marchés où sont utilisés les modèles qu'elle contribue à fonder. Le format d'un article interdit de reconstituer la totalité de cette histoire et d'explorer systématiquement ces deux scènes. Tout en résumant à grands traits, dans l'encadré 1, ce qui se joue sur la première de ces scènes (nous le décrivons plus systématiquement ailleurs : François et Frezal, 2016), nous concentrerons notre propos sur l'institutionnalisation de ces raisonnements en suivant leur pénétration dans un secteur particulier, celui des entreprises d'assurance. L'étude de la progression des mathématiques financières est particulièrement nourrie pour certains secteurs financiers comme la banque et les marchés financiers (voir par exemple MacKenzie et Millo, 2003) ; elle l'est cependant beaucoup moins pour le secteur assurantiel, traditionnellement moins étudié que les deux précédents (voir cependant Zelizer, 1979 et, plus récemment, Jarzabkowski *et al.*, 2015). L'un des enjeux de cet article est donc de contribuer à une sociologie du secteur de l'assurance, où le rôle des mathématiques financières est plus récent, où leur pénétration épouse des chemins spécifiques, et où leur application est à certains égards plus radicale encore que dans la banque ou dans les activités de marché. L'encadré 2 présente la façon dont nous avons mené l'enquête dans ce secteur.

Encadré 1. L'histoire chaotique d'une assimilation mathématique

L'histoire de l'assimilation des situations d'aléa et d'hétérogénéité se joue d'abord dans l'espace scientifique de la construction de la théorie des probabilités. Cette histoire, fort peu linéaire, se joue sur plusieurs scènes et à des époques qui se superposent partiellement. La première se

construit autour de la modélisation du *jugement individuel*, entre le milieu du XVII^e et la fin du XVIII^e siècle. L'enjeu est de modéliser le comportement d'un homme raisonnable, c'est-à-dire dégagé de la pollution de ses émotions, et de formaliser sa prise de décision. À mesure que s'édifie une construction théorique permettant de clarifier la notion d'espérance, il apparaît de plus en plus évident que les individus raisonnables ne se comportent pas en maximisant l'espérance. Cette première séquence conduit ainsi à une conclusion claire : l'espérance, quelle que soit la forme qui lui est donnée, et donc plus généralement tout outil statistique, ne permet pas d'éclairer les décisions à prendre en situation d'aléa.

D'autres arènes se saisissent des outils probabilistes *pour rendre compte de régularités collectives*. C'est en particulier en s'attachant à donner un sens à des données publiques qu'Adolphe Quételet va opérer un déplacement fondamental. Il constate en effet que la forme de la distribution des mesures successives que l'on effectue sur un même individu (une statue dont on mesurerait le torse) est la même que celle des mesures réalisées une et une seule fois sur un grand nombre d'individus (si, par exemple, l'on mesure le tour de poitrine de tous les membres d'un régiment). Selon l'expression consacrée, « tout se passe comme si » les caractéristiques physiques des individus réels se distribueraient autour de celles d'un « homme moyen », dont la moyenne des mesures serait un estimateur du tour de poitrine. A. Quételet impose ainsi l'idée que, puisque l'espérance est utile à qui fait face à l'hétérogénéité d'une distribution (l'intendance qui souhaite savoir quelle longueur totale de drap commander pour équiper l'ensemble du régiment), alors elle a aussi un sens pour un individu placé en situation d'aléa (par exemple un soldat qui, sans connaître les tailles des uniformes, est désormais fondé à répondre aveuglément qu'il taille M).

La troisième scène où va se jouer la naturalisation des assimilations dont A. Quételet a implicitement posé le principe se joue dans la mathématisation des probabilités. À l'issue d'un investissement collectif de plusieurs décennies, Andreï Kolmogorov propose en 1933 une axiomatisation dont l'abstraction a pour effet de soustraire à la discussion certains des principes qui y sont implicitement inclus : la théorie des probabilités s'appuie sur des fondements qui n'ont plus vocation à être discutés, en dépit des assimilations qu'elle peut comporter.

Encadré 2. Reconstituer une histoire implicite, ou la stratégie de l'oblique

Cet article s'appuie sur trois types de sources. Sur le modèle des travaux consacrés à la sociologie des outils de gestion, elle s'appuie d'abord sur une analyse systématique des outils financiers et prudentiels mobilisés par les compagnies d'assurance afin « d'ouvrir la boîte noire [qu'ils constituent], s'attachant à comprendre et décrire leur structure pour questionner l'évidence, [...] examiner et discuter les savoirs incorporés dans ces outils » (Chiapello et Gilbert, 2009, p. 328).

Elle mobilise ensuite les résultats d'une enquête par entretien menée auprès des acteurs du secteur de l'assurance (une trentaine d'entretiens semi-directifs ont été effectués avec des actuaires, des membres de l'Autorité de contrôle ou de la direction du Trésor, des membres de la direction des risques de grands groupes d'assurance, des consultants, etc.). Ces entretiens ne portaient pas sur la reconstitution de la trajectoire des raisonnements qui assimilent aléa et hétérogénéité : elle échappe en effet le plus souvent aux acteurs eux-mêmes, et les interroger frontalement sur la manière dont cette assimilation en est venue à jouer un rôle crucial dans leurs activités aurait entièrement surdéterminé la

situation d'entretien (Lizé, 2009). Les entretiens ont, le plus souvent, été réalisés autour de questionnements latéraux à la reconstitution de cette trajectoire. Deux types d'entretiens, plus précisément, ont été exploités. Les premiers portaient sur la reconstitution de la trajectoire de certains raisonnements financiers qui fonctionnent à maints égards comme des véhicules de l'assimilation qui constitue notre objet. Un second type d'entretiens alimente notre corpus : ceux que nous consacrons à la reconstitution des pratiques de certains professionnels de l'assurance qui mobilisent quotidiennement ces raisonnements financiers ou leurs résultats. Il s'agit en ce cas d'entretiens cherchant à rendre compte des décisions d'investissement, d'une part, et d'autre part d'entretiens portant sur le développement des activités de gestion des risques au sein des entreprises d'assurance.

Une troisième source, enfin, est constituée par un très gros volume de notes de terrain rassemblés dans le cadre d'une observation participante au sein de la direction des risques de la filiale française d'un grand groupe d'assurance européen effectuée entre 2011 et 2014. Mobilisée de longue date pour nourrir l'analyse du travail ouvrier (Peneff, 1996), l'observation participante est aussi souvent mobilisée dans les travaux portant sur les activités financières, en particulier lorsqu'elles engagent des dimensions techniques (Beunza et Stark, 2004). Ces observations permettent à la fois de repérer la circulation de certains raisonnements au cœur même des modèles de risque et de voir comment ces raisonnements (ou leurs résultats) étaient ensuite engagés dans les pratiques d'acteurs qui ne se réduisaient pas aux seuls membres de la direction des risques, mais aussi de la direction générale, de la direction financière et de la direction des investissements. Cette situation d'observation permettait de disposer d'un accès privilégié à des données qui échappent souvent à l'enquête. Les directions en question sont malaisément accessibles à un enquêteur qui resterait entièrement extérieur à l'entreprise (Diaz, 2005), et la

maîtrise technique des raisonnements permet par ailleurs d'en repérer plus facilement les traces et d'en reconstituer le destin. Ces notes ont fait ensuite l'objet d'une exploitation dont les principes sont très proches de ceux qui nous ont guidés dans l'exploitation des entretiens. Il n'y est que très rarement explicitement question de l'assimilation dont nous souhaitons retracer ici la trajectoire d'institutionnalisation, mais les situations de travail où cette assimilation est *de facto* en jeu sont en revanche très nombreuses. L'enjeu de l'exploitation était de repérer cette mise en jeu et d'en comprendre les principes et les modalités.

1. L'appropriation de l'amalgame par l'industrie financière

L'appropriation des raisonnements qui assimilent les situations d'aléa et d'hétérogénéité est un processus lent qui se joue d'abord dans les secteurs des activités de marché et de la banque, et qui ne pénètre le secteur assurantiel que tardivement et par contiguïté. Pour reconstituer cette lente appropriation de l'amalgame par la banque et la finance de marché, il faut distinguer les chemins qu'empruntent deux familles de modèles où l'amalgame est très inégalement à l'œuvre. Dans les premiers, centrés sur le calcul d'indicateurs de risque, l'amalgame est présent dès l'origine et va ensuite circuler, comme *incognito*, dans les pratiques de ceux qui l'utilisent, d'abord pour procéder à des décisions d'investissement, ensuite dans la mise en place de certaines règles prudentielles. Les modèles de la seconde famille, qui reposent sur le calcul stochastique, ne placent pas en leur cœur l'amalgame aléa/hétérogénéité. Ce sont les pratiques des acteurs qui, progressivement, vont l'y placer, en faisant jouer à ces modèles un rôle de plus en plus étendu. Ces deux chemins, longtemps distincts et indépendants l'un de l'autre, vont finalement converger, au cours des années 1990 — au moment même où ils vont être adoptés par les professionnels de l'assurance.

1.1. L'appropriation des indicateurs de risque, ou l'amalgame au cœur des pratiques

Si l'on suit la lecture qu'en propose Franck Jovanovic (2012), l'économie financière voit véritablement le jour au cours des années 1950 et 1960 de la rencontre de la théorie des probabilités et de la science économique. Dans les années 1950, des économistes, pour certains engagés dans l'étude des marchés financiers en mobilisant les outils de l'économétrie, s'appuient sur ce corpus désormais vulgarisé (Shafer et Vovk, 2001) pour donner à leur approche un tour plus théorique et plus systématique. Ils mobilisent alors des outils mathématiques permettant d'approcher les propriétés des variables aléatoires et des raisonnements en incertitude pour élaborer les premières propositions de l'économie financière.

C'est le cas en particulier de la théorie des choix de portefeuille de Harry Markowitz (1952) qui engage, de manière centrale, l'assimilation des situations d'aléa et d'hétérogénéité. Les méthodologies de choix de portefeuille commencent par évaluer, pour chaque portefeuille potentiel, les deux grandeurs statistiques qui forment son couple « risque/rendement ». L'enjeu est de déterminer la frontière d'efficacité représentant l'ensemble des portefeuilles qui soit maximisent le rendement pour un niveau de risque donné, soit minimisent le risque pour un rendement donné. Dans ces modèles, les raisonnements se placent dans un horizon temporel trop court pour que l'espérance puisse s'incarner dans une moyenne : la baisse de la valeur d'un actif en t ne pourra pas être compensée par une hausse en $t+1$. L'investisseur, autrement dit, se situe dans un horizon temporel qui lui interdit de mutualiser les hétérogénéités temporelles : il est en situation d'aléa. Dès lors, l'utilisation d'indicateurs statistiques relève de l'amalgame entre aléa et hétérogénéité. Si l'investisseur investissait à long terme, il pourrait négliger l'indicateur de risque, qui aurait disparu dans la loi des grands nombres ; mais s'il investit à court-terme, l'espérance de rendement n'est pas un estimateur du résultat qu'il compte obtenir. Les investisseurs qui s'appuient sur le couple

risque/rendement pour prendre leur décision pensent vivre dans un monde déterministe où l'indicateur de rendement serait un indicateur fiable du résultat, comme si le décideur était en situation d'hétérogénéité (Frezal, 2015b).

Lorsqu'elles furent publiées, ces théories étaient parfaitement ésotériques et réservées à une audience extrêmement réduite de *happy few* universitaires (Rubinstein, 2002). Ce n'est que très progressivement qu'elles s'imposent comme un référent cognitif pour les professionnels de la finance — et, dans cette conquête progressive, le rôle de la formation initiale fut évidemment d'une importance capitale. Aux États-Unis, elles sont enseignées à partir de la fin des années 1960 dans la plupart des *business schools*, où elles s'imposent comme un gage de scientificité et de rigueur dans un espace jusque-là orienté vers des formations très appliquées (Fourcade et Khurana, 2013).

Mais, pourra-t-on objecter, l'appropriation de l'assimilation aléa/hétérogénéité ne vaut à ce stade qu'au sein de certaines communautés et pour certaines activités, bien délimitées, et ces raisonnements n'affectent dès lors que certaines décisions : celles, en l'occurrence, qui président aux décisions d'investissement. Si l'amalgame des situations d'aléa et d'hétérogénéité pénètre aussi profondément le pilotage des organisations financières, c'est qu'on le retrouve au cœur des pratiques prudentielles qu'elles adoptent progressivement. C'est en particulier le cas dans le secteur bancaire, dont la régulation a connu au cours des années 1990 des inflexions décisives pour notre propos (voir Underhill et Zhang, 2008). Le premier système de règles arrêtées par le *Basel Committee of banking supervision*, connu sous le nom de Bâle I, repose sur un dispositif que l'on peut grossièrement qualifier de comptable, avant d'être économique et probabiliste : il pose que le ratio des fonds propres d'un établissement de crédit par rapport à son volume d'activité ne peut pas être inférieur à 8 % (ratio « Cooke »). Les accords de Bâle I vont très vite faire l'objet de vives critiques, portant en particulier sur l'accroissement du coût du crédit qu'il engendrerait. À la faveur notamment d'une mobilisation de certains acteurs privés (Underhill et Zhang, 2008, p. 543), un nouveau cadre prudentiel

commence d'être esquissé à la fin des années 1990. Pour leur partie quantitative, ces nouvelles règles reposent sur l'abandon d'un ratio global forfaitaire pour lui substituer un ratio fondé sur la quantification du risque associé à chaque ligne de crédit. Ce ratio est déterminé soit par l'utilisation d'une formule standard définie par les régulateurs, soit par le développement de mesures internes. À cette occasion, les régulateurs bancaires adoptent la notion de « *value-at-risk* » (VaR), qui désigne la probabilité que les pertes sur un portefeuille d'actifs excèdent un certain montant dans un horizon de temps donné. Dans les activités de marché, cette mesure statistique avait commencé à se substituer à la variance comme indicateur du risque d'un portefeuille lorsque J. P. Morgan avait tenté de l'imposer comme standard de marché *via* son système *Riskmetrics* (Linsmeier et Pearson, 2000), après son émergence à la fin des années 1980. Dans la continuité de la quantification introduite en finance par la théorie du portefeuille, l'amalgame entre aléa et hétérogénéité se retrouve ainsi au cœur de ces nouveaux dispositifs prudentiels qui font désormais reposer la gestion de l'aléa sur une grandeur statistique. Avec Bâle I, le « risque de crédit » désignait le fait que, sur les milliers ou millions de créances, une proportion donnée n'allait pas être remboursée. Les créances sur les débiteurs étaient corrigées d'une provision statistique correspondant à l'espérance de défaut appréciée sur la totalité du portefeuille de crédit : on était en situation d'hétérogénéité. Bâle II ajoute le nouveau mécanisme fondé sur la VaR, dont l'objectif est de fixer le montant minimal de fonds propres permettant d'absorber un choc en cas de dérive inattendue du taux de défaut moyen — en d'autres termes, en se plaçant dans une situation d'aléa.

À la fin des années 1990, l'amalgame aléa/hétérogénéité est désormais au cœur des pratiques de certains acteurs financiers. Logé au principe théorique de modèles développés au début des années 1950, il a d'abord structuré les pratiques des professionnels des marchés financiers après que ceux-ci, à partir de la fin des années 1960, y ont été formés dans les *business schools*. Au début des années 2000, à la faveur des vives polémiques qui se sont développées dans les

années 1990 autour des principes de régulation du secteur bancaire, l'amalgame élargit très sensiblement son emprise en se trouvant désormais au cœur des règles prudentielles qui s'imposent au secteur bancaire. C'est à partir de ces activités, contiguës comme nous le verrons aux métiers de l'assurance, que l'amalgame va y faire son entrée à partir de la fin des années 1990. C'est à cette époque que ce premier chemin qu'a emprunté l'amalgame pour se loger comme une institution au cœur des pratiques des acteurs financiers rencontre une autre voie, sensiblement différente, qu'il nous faut maintenant présenter.

1.2. Le calcul stochastique : des produits dérivés à la gestion prudentielle

Nous venons de le voir : les modèles de la théorie du portefeuille contiennent en leur cœur l'assimilation des situations d'aléa et d'hétérogénéité. Ce n'est pas le cas des outils stochastiques qui, à partir du début des années 1970, sont d'abord mobilisés pour valoriser les produits dérivés, avant d'être appropriés par des praticiens qui vont progressivement en étendre la sphère d'application et y réintroduire l'assimilation dont nous retraçons la trajectoire.

Au cours des années 1960, le calcul stochastique est mobilisé pour résoudre une difficulté qui, si elle a des implications pratiques, est d'abord comprise en un sens entièrement théorique : celle de la valorisation des produits dérivés. Un produit dérivé est un produit dont la valeur fluctue en fonction des variations de valeur d'un autre produit (une tonne de blé, une action, une obligation, etc.), que l'on désigne comme le sous-jacent. Il s'agit, par exemple, d'un « *call* », c'est-à-dire le droit d'acheter demain à un prix fixé aujourd'hui telle action ou telle quantité de blé, ou d'un « *future* », soit l'achat aujourd'hui, pour un prix donné, d'une action ou d'une tonne de blé qui me sera livrée à une date définie à l'avance et que je devrai alors régler. Ces produits dérivés, très anciens (en particulier sur les marchés agricoles où ils permettaient de se prémunir contre les crises frumentaires), ont longtemps été très délicats à tarifier : l'intuition, triviale, selon laquelle le prix du produit dérivé doit avoir quelque chose à voir avec celui du sous-jacent a longtemps été impossible à formaliser de manière

convaincante. Aussi les *traders* de Chicago qu'étudient Donald MacKenzie et Yuval Millo (2003) s'appuyaient-ils encore, au début des années 1970, sur des dispositifs conventionnels pour établir la valeur de ces produits. Or, en s'appuyant sur les outils du calcul stochastique, des économistes (de Chicago, également) vont définir une méthode, passée à la postérité sous le nom de « formule de Black-Scholes », qui propose une formalisation du lien entre le prix du produit dérivé et les caractéristiques de son sous-jacent.

Dans sa version théorique, la méthode de Black-Scholes n'implique pas que soient assimilées les situations d'aléa et d'hétérogénéité. Elle repose en effet sur l'idée que le cours du sous-jacent varie à chaque instant et qu'entre la date d'achat et la date d'exercice de l'option, un très grand nombre de variations seront observées. La formule de Black-Scholes propose de définir la valeur de l'option comme le coût cumulé des innombrables micro-rebalancements successifs de portefeuille qui permettent de se prémunir, à chaque instant, contre la variation infinitésimale du sous-jacent. Dans sa formalisation mathématique, le prix de l'option est l'intégrale d'un grand nombre de micro-coûts⁷⁶, dont chacun correspond aux micro-variations modélisées. Il s'agit donc bien de la description d'un phénomène d'hétérogénéité où les statistiques s'appliquent et, si le comportement statistique des variations de cours a été correctement modélisé, la valorisation de l'option est pertinente. De fait, en régime de croisière, les unités d'achat-vente de produits dérivés sont structurellement source de profits pour les banques — et ils le demeurent tant que les variations de cours sont correctement modélisées, ce qui cesse d'être le cas lorsque la volatilité augmente brusquement ou bien lorsque certaines hypothèses de liquidité ne sont plus vérifiées.

Sur un plan purement théorique, donc, la formule de Black-Scholes ne repose pas sur l'assimilation des situations d'aléa et d'hétérogénéité. Mais la portée de cette formule est immense, et elle va être engagée dans des pratiques qui introduisent, de manière centrale, cette assimilation. L'achat et la vente

⁷⁶ Ou, de façon équivalente, l'espérance d'un micro-coût qui lui est proportionnelle.

d'options sont conçus par les acteurs du marché comme une activité de « gestion du risque » financier, donc comme une manière de gérer les aléas des marchés. Dans les années 2000, Jean-Louis était analyste quantitatif dans une grande banque londonienne, et les modèles de son service n'étaient pas parvenus à anticiper l'irruption de la crise de 2007-2008. Après qu'elle est intervenue, ses supérieurs lui demandent, ainsi qu'à ses collègues, de tester différents modèles, reposant sur les outils tirés de la formule de Black-Scholes et qui auraient pu mieux la prévoir. Quelques années plus tard, en 2014, il constate son échec :

« Suite à 2008, il a fallu expliquer 2 milliards de pertes, c'est-à-dire trouver des modèles qui auraient permis de les éviter. On a testé plusieurs modèles plus complexes, de la vol sto [volatilité stochastique]... et en fait, ça ne marchait pas. On se rendait compte qu'on se serait fait exploser à chaque fois » (Entretien, mars 2014).

L'échec des équipes de Jean-Louis n'est pas fortuit. Les modèles qu'elles produisent sont capables de décrire l'hétérogénéité, mais les fluctuations de haute fréquence qu'ils permettent de décrire efficacement sont profondément modifiées en période de crise. Les investisseurs sont alors placés en situation d'aléa que les modèles, pertinents pour formaliser les situations d'hétérogénéité, ne peuvent par définition pas saisir. Le commanditaire de Jean-Louis assimile l'hétérogénéité des fluctuations *intraday* et l'aléa que constitue leur évolution brutale. Ils tentent de les capter par un même outil de « gestion des risques », les modèles d'évaluation des options, dont le périmètre de pertinence est censé s'être désormais considérablement étendu.

Au sein des mondes bancaires et des activités de marché, l'assimilation des situations d'aléa et d'hétérogénéité se fait au cours d'un lent processus d'appropriation de modèles mathématiques développés à partir des années 1950 qui, pour certains, contiennent en leur principe cette assimilation et qui, pour d'autres, sont utilisés de telle manière que cette assimilation se trouve au cœur de la pratique des acteurs qui les mobilisent. Dans les deux cas,

cependant, cet amalgame n'est pas reconnu comme tel : il s'installe au cœur des pratiques comme le passager clandestin de modèles mathématiques qui s'imposent parfois au cœur de certaines pratiques d'arbitrages localisés, comme sur les marchés financiers, et parfois au principe même des dispositifs prudentiels. C'est à partir de ces activités financières contiguës à l'industrie assurancielle qu'ils vont progressivement y pénétrer.

2. Appropriation et mobilisation de l'amalgame dans le secteur assurantiel

Entre la fin des années 1990 et le milieu des années 2000 toutefois, l'assimilation des situations d'aléa et d'hétérogénéité que se sont lentement appropriées les professionnels de la banque et du marché fait une entrée en force dans les pratiques des acteurs de l'assurance qui jusqu'ici l'ignoraient très largement. Dans un contexte marqué par une emprise croissante des logiques financières et par des opérations de fusion-acquisition d'envergure, les plus grandes entreprises d'assurance vont progressivement adopter des raisonnements et des outils qu'elles empruntent aux activités de marché ou aux grandes banques dont elles sont souvent très proches. Cette appropriation progressive va prendre un tour particulièrement radical, en particulier lorsqu'à cette phase d'appropriation souterraine et progressive succède une période de mobilisation des plus gros acteurs du marché assurantiel visant à contrôler les modalités de calcul de la valeur de leurs entreprises. À ce stade, cependant, la pénétration de l'amalgame aléa-hétérogénéité ne concerne encore que certains segments du secteur, dominants mais étroits : pour comprendre pourquoi ces raisonnements vont s'imposer à l'ensemble des acteurs — très hétérogènes — du monde assurantiel, la mobilisation de certains d'entre eux autour des nouvelles règles prudentielles, arrêtées en 2009 dans la directive Solvabilité II après avoir mûri au cours des années 2000, constitue un moment déterminant.

2.1. Transfert et appropriation

Les grands groupes d'assurance européens que nous connaissons aujourd'hui ont été constitués au cours de la seconde moitié des années 1990 par agrégation de nombreuses structures plus petites et jusqu'alors éparses. Pour ne citer que les opérations les plus importantes, on verra notamment se succéder en 1995 le rachat de La France puis sa fusion en 1996 avec La Concorde, constituant un ensemble qui servira de socle au futur groupe Generali France ; en 1996, le rachat de l'UAP par Axa ; en 1997, la fusion des AGF avec Athéna puis leur acquisition majoritaire par Allianz en 1998 ; en 1998 également, le rachat du Gan par Groupama, etc. Ces opérations de fusion-acquisition interviennent dans un contexte singulier. En France, elles coïncident avec la privatisation des plus grands groupes d'assurance (les AGF, l'UAP, le Gan) — c'est d'ailleurs la mise sur le marché de certains de ces groupes qui alimente la recomposition du secteur. Ces opérations se déroulent, par ailleurs, alors que l'emprise des logiques financières sur les plus grandes entreprises (en particulier lorsqu'elles sont cotées) ne cesse de s'affirmer, aux États-Unis d'abord (Davis *et al.*, 1994), mais aussi en Europe (François et Lemerrier, 2016). Cette emprise croissante des logiques financières se traduit notamment par une attention beaucoup plus soutenue à la communication destinée aux actionnaires (Rao et Sivakumar, 1999).

Dans le cas particulier des grands groupes d'assurance (Axa, Allianz, Generali, etc.), la communication financière s'appuie sur des outils qu'ils empruntent aux acteurs financiers présents de longue date sur la scène où ils pénètrent à cette époque. Chez Axa par exemple, les innovations en termes de gestion des risques apparaissent lorsque sont évoquées non les pratiques liées aux métiers de l'assurance, mais celles liées à des métiers adjacents portés par des organisations partenaires : c'est le cas par exemple de la *Value-At-Risk*, qui est évoquée dès le rapport annuel 1997 dans le chapitre portant sur « Gestion des risques et sensibilité aux risques de marché » pour rendre compte de la mesure du risque dans l'entreprise d'*Asset Management* Donaldson, Lufkin et Jenrette

(DLJ), que contrôle Axa — et c'est à cette occasion que la VaR est présentée au lecteur :

« La mesure de l'exposition aux risques de marché utilisée par DLJ est la "Value At Risk" ("VAR"). [...] Cette mesure tient compte des corrélations entre les différents types de risques de marché auxquels est soumis DLJ. Cette VAR peut être calculée par type de risques sans tenir compte des corrélations pouvant exister entre les différents risques. » (Axa, *Rapport annuel* 1997, p. 99).

La gestion des risques est importée, au sein des grands groupes d'assurance, par des entreprises financières qui leur sont proches (chez Allianz, c'est la Dresdner Bank, achetée par Allianz au tout début des années 2000, qui joue un rôle déterminant) mais qui interviennent sur des segments d'activité où les nouvelles techniques prudentielles ont été adoptées plus tôt — et dans cette importation circulent aussi certains raisonnements congelés dans des outils financiers, et dont ils constituent des rouages fondamentaux.

Un autre mécanisme contribue à faire pénétrer au cœur des entreprises d'assurance ces outils financiers et les amalgames qu'ils peuvent impliquer : la soumission aux standards imposés par les agences de notation. Dès 2000, Allianz souligne que sa gestion des risques est soumise à l'évaluation des autorités de supervision et des agences de notation — et que, dans ces conditions, le groupe dispose des meilleures notes possibles (Allianz, *Annual report* 2000, p. 133). Dans le rapport 2002, Allianz annonce être en train de développer des modèles internes — et, dans l'intervalle, de s'en remettre aux modèles de Standard and Poor's :

« In 2002, we launched a comprehensive project which substantially improved internal risk analysis in the insurance business. (...) By 2004, we expect to be able to convert our value-oriented controlling to risk capital, that was calculated with help of this model. At present, we are using risk capital models provided by the Standard & Poor's rating agency » (Allianz, *Annual report* 2002, p. 73).

Les modèles de Standard & Poor's ont une sorte de fonction matricielle, qui se comprend d'abord en ce que, comme les grandes banques ou les sociétés d'*asset management*, les agences de notation ont adopté relativement tôt les nouvelles métriques de risque, plus tôt en tout cas que les grands groupes d'assurance.

Les premières voies de circulation des outils cognitifs où se loge l'amalgame aléa/hétérogénéité relèvent ainsi de ce que l'on pourrait appeler une appropriation par contiguïté : soumises à de nouvelles exigences en matière de communication financière — les plus grands groupes d'assurance sont désormais privatisés et cotés, et ils sont engagés dans des opérations de fusion-acquisition délicates dans un contexte où le pouvoir des investisseurs ne cesse de croître —, les entreprises d'assurance empruntent à des acteurs (banques, fonds d'investissement) dont elles sont proches les outils financiers que ceux-ci ont assimilés, et dont l'amalgame constitue une pièce maîtresse. À ce premier mécanisme d'appropriation implicite et souterrain va rapidement s'articuler une mobilisation délibérée, qui va en démultiplier la portée.

2.2. Évaluer les entreprises d'assurance : la mobilisation des CFO

À la fin des années 1990, parce qu'ils se soumettent aux jugements des analystes financiers et aux verdicts des agences de notation, les grands groupes d'assurance européens s'attachent à contrôler la nature des indicateurs qui sont produits pour rendre compte de leur santé financière. C'est dans ce cadre que les directeurs financiers (*chief financial officers*, CFO) des principaux groupes européens, réunis au sein d'une sorte d'association professionnelle, le CFO Forum, tentent de normaliser les méthodes d'évaluation des portefeuilles d'assurance (et donc de valorisation des entreprises), afin de proposer aux analystes financiers un outil de comparaison d'entreprises implantées sur des marchés différents et dont les contrats sont hétérogènes. Ils vont pour cela utiliser les outils de valorisation de portefeuilles qu'ils se sont appropriés. En 2004 est publiée l'*European Embedded Value* (EEV, CFO Forum, 2004), qui permet d'estimer la valeur d'un portefeuille d'assurance-vie en standardisant

l'actualisation des *cashflows* futurs. Ces outils, placés au cœur de la communication des entreprises, font par ailleurs partie des outils de pilotage des entreprises elles-mêmes : les décisions de lancement d'un nouveau produit ne s'appuient plus tant sur les anciens concepts de profit et de rendement que sur la *New Business Value*, autrement dit sur l'*embedded value* créée par les nouvelles affaires, et sur la *New Business Margin*, à savoir son montant rapporté au volume d'affaires nouvelles. En interne comme en externe, les outils financiers, appuyés sur les modèles issus de la théorie du portefeuille, gagnent donc du terrain.

Ces outils vont rapidement être amendés pour mieux tenir compte des spécificités des entreprises d'assurance — et ces amendements vont démultiplier l'emprise de l'amalgame des situations d'aléa et d'hétérogénéité, en élargissant encore l'emprise des outils issus du calcul stochastique. L'*European Embedded Value* se contente en effet de projeter les perspectives de *cash flows* futurs dans un scénario central d'évolution des marchés financiers. Or, les contrats d'assurance-vie offrent un profil de *cash flows* dissymétrique : lorsque les marchés financiers sont favorables, les profits sont partagés entre assureur et assurés, mais lorsqu'ils évoluent de façon défavorable, les pertes sont à la charge de l'assureur (Frezal, 2016). Cette dissymétrie notamment les apparente à des options financières : pour une option d'achat, par exemple, toute hausse du sous-jacent au-dessus du prix d'exercice de l'option se traduira par un gain alors qu'une baisse en deçà ne se traduira pas par une perte. Lorsque l'EEV valorise ces contrats en les projetant sur un scénario central, elle ne capte pas ces non-linéarités et peut dès lors conduire à valoriser deux contrats à un niveau comparable alors que leurs comportements, en cas de scénario financier adverse, seraient très différents — par exemple parce que les seuils de déclenchement des options sont tous deux en deçà du scénario central mais ne sont pas identiques. Les groupes de travail du CFO Forum ont donc proposé de valoriser le coût de ces options pour l'assureur. Cette valorisation a été confiée aux actuaires qui étaient alors chargés des calculs d'EEV et qui avaient

été au contact, en particulier durant leur formation, des outils de *pricing* des options financières (*puts*, *calls*, etc.) développés par la finance de marché ; ils avaient d'ailleurs déjà commencé d'appliquer ces outils à l'assurance (voir par exemple le mémoire d'actuariat de Merlus et Pecqueux, 2000). Ils vont donc facilement les importer et développer des modèles de valorisation stochastique des garanties, options explicites ou implicites incorporées dans les contrats d'assurance-vie.

L'EEV est dès lors complexifiée, et le CFO Forum lui substitue la *Market Consistent Embedded Value* (MCEV, CFO Forum, 2008), qui entérine pour le secteur des assurances cette fusion des outils de finance d'entreprise avec les technologies de finance de marché. Désormais, au lieu de projeter les *cash-flows* futurs dans un scénario financier unique, dit « central », une nouvelle convention s'impose. Face à l'aléa des marchés financiers, on élabore des jeux de scénarios financiers simulant l'évolution d'un grand nombre de paramètres sur plusieurs années ou plusieurs décennies (niveau des taux, des indices action ou des indices immobilier, etc.). Pour chaque scénario, les modèles évaluent la chronique des *cash-flows* qui, actualisée, fournit la valeur actuelle des *cash-flows* associée au scénario considéré. La MCEV correspond alors à la moyenne sur l'ensemble des scénarios simulés des valeurs ainsi déterminées. Les coûts d'option correspondent ainsi à l'écart entre la valeur moyenne du portefeuille sur l'ensemble des scénarios simulés et la valeur du portefeuille projeté sur le scénario central. Alors qu'une unique trajectoire financière se réalisera, on se réfère à l'espérance des gains futurs : les valeurs des portefeuilles d'assurance reposent sur un indicateur statistique alors qu'il s'agit d'une situation d'aléa.

Si ces outils sont issus des techniques de valorisation des options conçues par les mathématiques financières, les options qu'on retrouve dans les contrats d'assurance ne sont pas de même nature, et elles placent en leur cœur l'assimilation des situations d'aléa et d'hétérogénéité. Le prix d'une option financière, rappelons-le, est calculé en construisant un mécanisme de couverture répliquant les variations à haute fréquence du sous-jacent et dont le

coût cumulé est statistiquement déterministe. Les assureurs transposent les méthodes de simulation à une situation très différente, puisqu'ils projettent leur bilan sur plusieurs années, prennent une photographie de la valeur finale de ce bilan, et fondent le coût d'option sur une moyenne des photographies alternatives envisagées. Ce changement d'horizon temporel — le micro-instant séparant deux rebalancements de portefeuille *versus* la durée d'un passif d'assurance de plusieurs années — transforme totalement la nature du phénomène considéré — une micro-variation qui va se renouveler *versus* une valeur finale qui sera observée une fois et une seule.

En répliquant les technologies de valorisation d'option issues des marchés financiers pour produire une estimation des options implicites contenues dans leurs portefeuilles d'assurance-vie, les compagnies d'assurance ont appliqué les outils statistiques développés pour la caractérisation de situations d'hétérogénéité à des situations d'aléa. Les statistiques comme outil d'appréhension, de synthèse et de gestion des aléas financiers sont désormais placées au cœur du pilotage des compagnies d'assurance-vie ; elles alimentent ainsi les décisions de mesure de la rentabilité, de lancement de produit ou de choix d'allocation stratégique d'actifs.

2.3. Définir de nouvelles règles du jeu

L'appropriation progressive et souterraine des modèles mathématiques, la mobilisation autour de certaines techniques de valorisation et de pilotage des entreprises ne concernent cependant qu'un nombre très limité de compagnies ou de mutuelles : celles qui sont engagées dans des opérations de fusion ou dont la taille est telle qu'elle les implique dans les réflexions du CFO Forum, celles enfin qui, plus généralement, sont engagées dans une démarche générale d'optimisation financière de leurs activités. L'amalgame aléa/hétérogénéité est bien *a priori* (placé au cœur des modèles de valorisation, il est d'autant moins susceptible d'être remis en cause qu'il est très largement ignoré) et normatif (les modèles qui engagent cet amalgame jouent un rôle déterminant dans le pilotage

des risques des plus grands groupes d'assurance), mais il est encore loin d'être *partagé* par tous les acteurs. En effet, le secteur assurantiel est extrêmement hétérogène et les très petits acteurs, en particulier, y sont extrêmement nombreux : fin 2007, après une décennie de consolidation du secteur, on dénombrait encore plus de 1 500 organismes d'assurance en France (ACAM, 2008, p. 6). Ces très petits acteurs n'ont *a priori* que de très faibles incitations à adopter ces nouvelles formes de raisonnement, beaucoup plus techniques que celles qu'ils utilisent dans leur gestion depuis des décennies. À l'image de ce que décrivent Céline Baud et Ève Chiapello (2015) dans le secteur bancaire, l'un des principaux véhicules de la généralisation des modèles financiers dans le secteur assurantiel est le changement spectaculaire de réglementation prudentielle intervenu à la fin des années 1990. Le cadre prudentiel qui prévaut alors est le fruit d'une lente élaboration qui court de la fin des années 1940 à la fin des années 1990, durant laquelle la conception qui sous-tend les dispositifs consiste à dégager un ou plusieurs ratios appelés à fonctionner comme un signal d'alarme. Ces ratios doivent être aisément compréhensibles par l'ensemble des acteurs du marché — y compris par les consommateurs — et être adaptables aux diverses situations nationales.

L'abandon de ces règles prudentielles et la mise au point, puis l'adoption et l'entrée en vigueur, d'un nouveau dispositif désormais connu sous le nom de Solvabilité II marquent le passage d'un monde où la réglementation distinguait *de facto* les situations d'aléa des situations d'hétérogénéité à un système où les deux types de situations sont assimilés. D'un point de vue quantitatif, l'appréciation de la solvabilité et des capacités de résistance d'une entreprise d'assurance reposent sur deux piliers : d'abord, l'évaluation du bilan, qui permet d'apprécier le montant de ses passifs (ses engagements), de ses actifs (ses investissements), et de les mettre en regard pour s'assurer que ses fonds propres (sa richesse nette) sont positifs ; ensuite, la fixation des exigences de capital, c'est-à-dire d'un matelas de sécurité suffisamment épais pour absorber les chocs imprévus, seuil en dessous duquel ses fonds propres ne doivent

normalement pas descendre. Ces deux piliers ont été profondément refondus par Solvabilité II, et le nouveau système place en son cœur les différentes déclinaisons de l'amalgame aléa/hétérogénéité.

On le voit, par exemple, dans l'évaluation des fonds propres. Dans l'ancien système, les fonds propres correspondaient au niveau de richesse accumulée jusqu'à la date d'inventaire et, si l'on risquait en raisonnant ainsi de pécher par optimisme, la contingence était gérée par la prudence. Avec Solvabilité II, ils prennent en compte les perspectives de profits futurs sur plusieurs décennies — profits futurs qui sont contingents à l'évolution des marchés financiers. Les principes de la MCEV sont alors importés dans le bilan prudentiel : l'espérance de ces profits futurs, devenue l'estimateur de ce qu'on va encaisser ou décaisser, s'assimile désormais à la richesse actuelle de l'entreprise. Ces nouvelles règles prudentielles qui fondent l'évaluation des fonds propres sur une statistique synthétique assimilent de fait deux positions : celle de l'actionnaire qui diversifie ses placements dans un grand nombre d'entreprises et qui est placé en situation d'hétérogénéité, et celle du dirigeant qui tente d'évaluer la capacité de résistance de l'entreprise qu'il dirige, et qui est en situation d'aléa puisque son entreprise sera confrontée à un unique scénario financier. Plus largement, cet amalgame se retrouve dans les principales orientations techniques de Solvabilité II, notamment pour l'évaluation des passifs et la définition de l'exigence de capital (Frezal, 2015a).

Nous explorons en détail, ailleurs, l'enchaînement de causes qui fonde l'abandon de l'ancien système et l'adoption du nouveau (François, 2015). Depuis la fin des années 1950 se font jour, lentement et progressivement, des mutations majeures dans l'appréhension des catégories financières et comptables : le développement de la théorie financière dont nous parlions plus haut, la remise à plat des normes comptables avec l'adoption des normes IFRS et, plus fondamentalement, de la comptabilité dite en *fair value* (Chiapello, 2005) et la redéfinition des normes prudentielles bancaires dans le cadre de Bâle II et III (Underhill et Zhang, 2008). Au cours des années 1990, ces trois mutations

initialement indépendantes convergent et font système : les hypothèses et les modèles de la théorie financière occupent ainsi une position centrale dans les nouvelles normes comptables et bancaires.

Ces cadres cognitifs, très largement étrangers au monde de l'assurance, vont alors être placés au cœur des nouvelles règles prudentielles. La cheville ouvrière de ce déplacement est l'entité « assurance » de la Commission européenne, qui annonce en 1999 qu'elle va entièrement refondre les règles prudentielles longuement mises au point depuis la Seconde Guerre mondiale. Ce revirement subit, rendu possible par le renouvellement partiel de cette petite structure (elle comporte une dizaine de personnes) relativement autonome, est porté par une forme de recherche de cohérence. Les règles comptables obéissent désormais aux principes de la *fair value* : la valeur des actifs est estimée en « valeur de marché » (au temps t , la valeur que l'on impute au bilan de l'entreprise correspond à ce que vaudrait cet actif si on le vendait sur un marché en cet instant précis), tandis que jusque-là la valorisation des actifs se faisait en « valeur historique », c'est-à-dire en imputant au bilan le prix d'acquisition de cet actif (Chiapello, 2005). Les membres de l'entité assurance voient comme une forme de logique à ce que les règles prudentielles soient elles aussi davantage « économiques ». C'est aussi le cas des règles bancaires : si l'on veut pouvoir contrôler des conglomérats financiers — des organisations qui proposent simultanément des services bancaires et des services d'assurance —, il devient nécessaire de construire une régulation point trop éloignée, ou au moins compatible, avec celle retenue pour les banques.

Mais cette volonté de mise en cohérence n'est pas seule en cause. Certains acteurs du marché défendent le principe d'une refonte d'ensemble du système, et vont faire pression en ce sens. C'est le cas de certains régulateurs, en particulier des régulateurs britanniques et néerlandais. Qu'ils soient motivés par une adhésion aux principes fondamentaux de la régulation prudentielle que nous présentons plus haut, ou par les intérêts de la place dont ils émanent (le *reporting* systématique qu'impose Solvabilité II intéresse par exemple les gros

cabinets d'audit qui peuplent la place de Londres), certains régulateurs, sans pouvoir imposer unilatéralement les options qu'ils souhaitent voir privilégier, vont s'attacher à peser en leur faveur. Alors que les options fondamentales de ce qui deviendra Solvabilité II sont encore en discussion, au tout début des années 2000, ils adoptent de nouvelles règles pour leur propre marché et endossent des options particulièrement radicales en matière de recours aux modèles financiers. De même, les plus grands groupes d'assurance européens, qui ont fait mouvement plus tôt que leurs concurrents vers ces nouveaux cadres cognitifs, et qui par ailleurs disposent, pour les assimiler, de moyens sans commune mesure avec des entreprises de taille souvent très modeste, y voient un moyen de pousser leur avantage.

Henri de Castries, président d'Axa, explique ainsi sans excès d'euphémisation que « Le bénéfice de la diversification risque d'entraîner un désavantage concurrentiel pour les acteurs moins sophistiqués ? Oui, et alors ? » (H. de Castries, *Les Échos*, 23 juin 2008).

Cette mobilisation des plus grands groupes d'assurance épouse deux voies principales. Chacun intervient d'abord directement auprès de l'entité assurance de la Commission et, plus généralement, de l'ensemble des acteurs engagés dans la négociation des nouvelles règles du jeu (et notamment des Trésors nationaux). Mais ces principaux groupes se mobilisent aussi en intégrant systématiquement les associations professionnelles et les instances de coordination inter-organisationnelles — comme le CFO Forum qui, par définition, ne comprend que des CFO de grands groupes précocement engagés dans ces transformations, ou, en France, la Fédération française des sociétés d'assurance (FFSA), à qui ils fournissent une expertise que les plus petites entités ne sont pas en mesure de contredire. Ces instances sont elles aussi reçues à Bruxelles ou au Trésor, et elles aussi contribuent, de manière plus (pour le CFO Forum) ou moins (pour la FFSA) militante, à défendre les options que retiendra finalement Solvabilité II.

Face à cette mobilisation de quelques acteurs dominants qui poussent à la formulation de nouvelles règles du jeu, l'opposition ne se structure qu'imparfaitement. Ceux des superviseurs qui ne militent pas pour l'adoption de Solvabilité II ne s'y opposent pas non plus de manière délibérée. Le superviseur français, par exemple, est assez profondément divisé entre certains contrôleurs qui voient dans Solvabilité 2 une mutation aventureuse et d'autres qui y sont *a priori* plus favorables. Confrontée à des réformes d'envergure qui s'achèveront par la fusion (fort mal vécue du côté assurance) avec le superviseur bancaire, l'autorité de contrôle des assurances ne prend qu'assez tard la mesure des implications — que ses membres jugent de manière par ailleurs ambivalente — de Solvabilité II. Quant aux autorités de contrôle allemandes, elles endossent très tôt les nouveaux principes qui, au début des années 2000, commencent à se dessiner (elles proposent dès 2002 une nouvelle approche dite « *risk-based* », proche dans son inspiration de ce qui commence d'être évoqué pour Solvabilité II) et, dans les négociations ultérieures, se concentrent non sur la philosophie générale du système mais sur des points identifiés comme très stratégiques pour les grands groupes allemands. De la même manière, les petites entreprises d'assurance européennes, qui ont beaucoup à perdre avec ces nouvelles règles qu'elles ne pourront que très difficilement assimiler, ne réagissent que tardivement : très nombreuses, de taille souvent très modeste, éloignées des centres de pouvoir, dotées de compétences faibles ou inexistantes en matière de mathématiques financières, beaucoup ne prendront la mesure des changements à venir que lorsqu'ils seront inéluctables. Sans doute parviennent-elles à contredire certaines orientations lorsqu'elles leur sont trop évidemment défavorables, comme lorsque les mutuelles françaises obtiennent, en 2006, que la faiblesse structurelle de leur capitalisation ne les handicape pas trop fortement. Mais ces aménagements ne leur permettent pas de lutter contre l'adoption de principes fondamentaux que les plus grands groupes ont largement anticipés.

L'adoption de Solvabilité II, au sortir d'une lutte dont on mesure le déséquilibre, constitue une étape déterminante dans le transfert des outils issus des mathématiques financières au secteur assurantiel : ses principes vont en effet s'imposer à l'ensemble des acteurs du secteur. Dans un secteur où subsistent de très nombreuses très petites entreprises, la généralisation de ces outils passe avant tout par la mise en place de nouvelles règles que les entreprises ne peuvent refuser d'appliquer.

3. L'appropriation reconduite de l'amalgame

Si l'on suit les propositions théoriques que nous précisons en introduction, une règle, une catégorie de pensée ou un raisonnement ne sont des institutions — c'est-à-dire des règles, des catégories ou des raisonnements partagés, normatifs et *a priori* — qu'à la condition (et donc, aussi longtemps) qu'on les utilise comme tels. Après les deux épisodes de mobilisation, autour de la MCEV et de Solvabilité II, qui ont contribué à faire pénétrer cet amalgame dans le secteur assurantiel, comment les acteurs des entreprises se le sont-ils approprié ? La mobilisation de cet amalgame peut parfois se faire sur un mode routinier et très peu réflexif, et aller jusqu'à relever d'une adhésion inconditionnelle ; elle peut aussi faire l'objet d'une mise à distance réflexive qui, dans certains cas, peut épouser les formes d'une manipulation cynique.

3.1. Le registre de la foi

La présence des modèles mathématiques dans les pratiques des acteurs financiers a quelque chose de paradoxal. Leur mise au point et leur mobilisation se justifient souvent dans le registre de la rationalité scientifique : combinant le double privilège de la rigueur apparente et de l'ésotérisme manifeste, ils sont censés substituer aux pratiques routinières de nouvelles logiques entièrement rationalisées. Et pourtant, lorsqu'ils sont évoqués par les acteurs qui les produisent ou les utilisent, le registre lexical est souvent celui de la foi. Y croire ou ne pas y croire, telle semble bien être la question — et

certains acteurs, de ce point de vue, sont des fidèles qui peuvent être blessés lorsqu'ils croisent la route d'un agnostique... Pauline travaille dans un grand groupe d'assurance européen, et elle fait partie des équipes qui, depuis plusieurs années, travaillent au développement des modèles internes de ce groupe. Elle raconte l'interaction un peu acide qui s'est déroulée lors d'un cocktail donné après que le modèle a été validé par le régulateur — étape décisive et redoutée dans la mise en conformité avec les règles de Solvabilité II (« S2 ») :

« Notre PDG nous a déclaré au pot S2 qu'il nous félicitait pour [l'approbation du modèle interne], mais que le modèle interne il y croyait pas quoi, "mais c'est bien que vous, vous y croyiez". [...] Ce qu'il voulait dire par là c'est qu'un modèle c'est un modèle, avec ses limites, et qu'il faut l'utiliser comme tel quoi, en étant moins radical. Ça a fait un peu crisser des dents. "Tant que j'ai pas passé une bonne crise pour voir si le modèle tient, j'y croirai pas" » (Entretien, octobre 2016).

Le différentiel de posture qu'entretiennent, à l'égard du modèle interne, le PDG et les équipes chargées de le mettre au point signalent tout ce qui sépare une hypothèse, que l'on adopte provisoirement mais qui est susceptible d'être remise en cause (« *Tant que j'ai pas passé une bonne crise pour voir si le modèle tient, j'y croirai pas* »), et l'institution, à laquelle l'adhésion peut être inconditionnelle (« *c'est bien que vous, vous y croyiez* »). C'est ce type d'adhésion inconditionnelle que l'on rencontre chez nombre d'utilisateurs ou de producteurs des modèles, qui rappellent de loin en loin, dans les discussions informelles avec leurs collègues, leur foi dans les principes qui fondent leur activité quotidienne. Richard, par exemple, est le *principal* d'un cabinet de conseil en actuariat. De nationalité australienne il est spécialiste des modèles de calcul de la MCEV et des exigences de capital Solvabilité II. Il justifie ainsi en quoi s'intéresser à la moyenne de mille scénarios projetés aléatoirement plutôt qu'à un autre indicateur tel que la médiane est préférable : « La moyenne permet de tout capter : si on change un seul scénario, la moyenne bouge, donc on capte tout » (Journal de terrain, janvier 2014).

Mais l'adhésion inconditionnelle qui peut présider à la mobilisation d'une institution n'est sans doute jamais mieux mise en évidence que lorsqu'elle est remise en cause — et l'on voit alors qu'elle peut concerner les modèles pris dans leur ensemble, mais aussi certains des raisonnements qui y sont engagés, comme l'amalgame qui ici nous occupe. La discussion animée qui se fait jour lors d'une réunion préparatoire au comité d'investissement de la filiale française d'un groupe européen en est un bon exemple. L'enjeu de la discussion au sein de la petite dizaine de participants cadres supérieurs ou dirigeants occupant des fonctions de pilotage (finances, investissements, risques centralisés, ou pilotage d'une *business unit* d'assurance-vie) concerne le montant de trésorerie à conserver en portefeuille. Le dilemme est très classique : d'un côté, il est préférable d'avoir le moins de trésorerie possible, puisque cette classe d'actifs rapporte peu ; mais il faut, d'un autre côté, en avoir suffisamment pour être en mesure de tenir face à une vague de retraits d'épargne massifs, sauf à devoir alors réaliser des actifs en urgence et partant prendre le risque d'essuyer des pertes. Tout l'enjeu est donc de placer le curseur entre ces deux injonctions contradictoires. En l'occurrence, dans un contexte de crise financière latente (interrogations sur la résistance du secteur en cas de défaut sur la dette grecque, rumeurs sur des changements de politique fiscale, etc.), l'entreprise craint d'avoir à faire face à une vague de retraits. Si tel devait être le cas, elle doit pouvoir disposer d'une trésorerie de 10 milliards d'euros. Si, au contraire, il n'y a pas de vague de panique et que seuls les rachats habituels interviennent, alors les besoins de trésorerie sont très faibles, car la collecte usuelle et les tombées de coupons suffiront à y faire face. L'un des participants à la réunion est un cadre de direction, formé initialement comme ingénieur et comme actuaire. Il explique : « nous avons une chance sur deux de devoir sortir 10 milliards, une chance sur deux rien [zéro euro, en absence de panique], donc mettons de côté cinq milliards ». Une discussion animée s'engage avec un autre participant, qui tente de le convaincre que cette option, parce qu'elle conduit à « perdre du rendement si tout va bien et ne pas être en mesure de résister en cas de

panique », est la pire de toutes celles que l'on peut envisager. Après un quart d'heure d'échange avec l'assemblée, ce dernier parvient à convaincre son collègue que le choix d'une position statistique, « en espérance », est en l'occurrence inadapté. Spontanément, le raisonnement statistique lui semblait pertinent pour décider face à cette situation d'aléa.

L'adhésion inconditionnelle est d'autant plus susceptible de se faire jour que sont mis en place des dispositifs qui mettent la réalité à distance, et qui donc soustraient — non par accident, mais par principe — les raisonnements au verdict des faits. Le « raisonnement en risque neutre » constitue l'un de ces dispositifs. L'objectif des modèles de valorisation des options est, nous l'avons vu, de lier la valeur d'un produit dérivé à celle de l'actif sous-jacent, que l'on doit par conséquent pouvoir modéliser. Cette modélisation repose, *a priori*, sur des calibrages empiriques : on raisonne alors en *probabilité historique*, ou en *monde réel*. Pour simplifier des calculs extrêmement lourds a été développé un artefact mathématique, très couramment utilisé sur les marchés financiers : la probabilité *risque neutre*. Est alors défini un univers virtuel, où les probabilités d'évolution des actifs ont été modifiées de telle sorte que tous aient la même espérance de rendement, celui d'un actif non risqué, et qu'ils ne diffèrent que par leur volatilité, plus ou moins élevée. Sous certaines hypothèses, on démontre alors mathématiquement qu'il est équivalent de valoriser les actifs *sous probabilité risque neutre* ou bien *en monde réel*. Suivant cette pratique usuelle sur les marchés financiers, la MCEV définie par le CFO Forum et Solvabilité II impose *de facto* de fonder des calculs sur un jeu de scénarios financiers *risque neutre*, et non *monde réel*. Cette transposition au secteur de l'assurance comporte des faiblesses théoriques intrinsèques (le fait, notamment, de faire interagir des actifs *risque neutre* avec des engagements *monde réel*), mais pour notre propos l'important est ailleurs : il est d'autant plus facile de faire jouer aux raisonnements un rôle *a priori* que les dispositifs qui pourraient participer à leur remise en cause sont mis à distance, et notamment l'épreuve des faits. Max est associé dans un cabinet de conseil spécialisé en actuariat. Il évoque ainsi le

hiatus qui se fait jour entre la perception du client, qui intuitivement considère que le jeu de scénario mobilisé est invraisemblable, et la méthode, qui rend la disqualification impossible :

« Lorsque nous avons fourni un jeu de scénarios financiers pour qu'ils puissent faire les calculs, son dirigeant m'a dit "40 % de scénarios où les taux sont négatifs, je n'y crois pas". Mais je lui ai répondu qu'il n'y avait pas à y croire ou non : l'intuition n'a aucune importance, le risque neutre n'a pas à avoir de lien avec la réalité ! » (Entretien, novembre 2016).

L'adhésion inconditionnelle à certains raisonnements relève ainsi de motifs contrastés, pour ne pas dire contradictoires. Y est d'abord à l'œuvre la reconduction d'une forme d'adhésion scientifique, particulièrement présente chez certains professionnels qui, comme les actuaires, jouent un rôle déterminant dans les entreprises d'assurance. L'outillage mathématique des raisonnements qu'ils mettent en œuvre, acquis durant leur formation initiale, fait l'objet d'une adhésion d'autant moins susceptible d'être remise en cause qu'elle constitue la pierre de touche de leur identité professionnelle et qu'elle est la ressource déterminante qu'ils peuvent mobiliser dans les luttes de juridiction qui les opposent à d'autres professions. Dès lors que cet outillage importe, comme un passager clandestin, l'amalgame entre aléa et hétérogénéité, celui-ci non plus ne sera pas remis en cause. Mais dans cette adhésion inconditionnelle se joue également, pour une bonne part, une forme de routine : la « mise en boîte noire » qu'opèrent les raisonnements scientifiques a été de longue date soulignée par la sociologie des sciences (Latour et Woolgar, 1988) ; elle a pour effet de rendre certains raisonnements inquestionnables et de produire des comportements routiniers qui s'appuient sur eux en les considérant, de fait, comme des points acquis.

3.2. Réflexivité et cynisme dans l'usage des institutions

La foi dans certains raisonnements n'est que l'un des modes d'usage de l'amalgame qui contribue à en faire une institution. Une fois placés au cœur de

cette activité, ces modèles offrent à ceux qui peuvent s'y référer des avantages fonctionnels qui contribuent eux aussi à expliquer leur emprise : en dépit de leurs limites intrinsèques ou des risques dont leur usage est porteur, ils rendent des services que les acteurs jugent supérieurs aux difficultés qu'ils soulèvent.

L'amalgame entre aléa et hétérogénéité subsume l'appréhension d'une alternative en un chiffre. L'indicateur statistique substitue ainsi à la contingence un résultat : l'immense vertu de ces outils est de rendre les situations commensurables et, en autorisant le calcul, de permettre de conclure. Cette supériorité rhétorique du chiffre est une vertu associée de longue date à la commensuration statistique lorsqu'on l'engage dans des disputes (voir à cet égard Espeland et Stevens, 1998, p. 416-422). Les acteurs financiers justifient ainsi l'usage des outils statistiques pour décider en aléa : ne pas disposer de chiffres les laisserait profondément démunis, et l'utilisation des modèles statistiques pour appréhender des situations d'aléa leur en propose. Alain est le CEO (*chief executive officer*) d'un organisme d'assurance qui entretient avec la modélisation une relation de scepticisme : il n'y a pas été formé lors de ses études, et il a rejoint le secteur de l'assurance assez tard dans sa carrière, pour y occuper des fonctions stratégiques. Confronté lors d'un échange informel à l'argument selon lequel les situations d'aléa ne se laissent pas saisir à l'aide des outils statistiques, il réagit :

« Certes, mais sans outils de type Black-Scholes par exemple, on ne saurait pas tarifier une option. Or, même si c'est faux, on a besoin d'une estimation pour décider » (Journal de terrain, Janvier 2016).

C'est le même type d'argument que mobilise Marc, le directeur de l'actuariat d'un petit groupe d'assurance, ingénieur et actuaire de formation, qui a débuté sa carrière à l'Autorité de contrôle au moment où Solvabilité II se mettait progressivement en place, avant de rejoindre plusieurs assureurs. Lors d'un entretien, lorsqu'il explique que le recours à des outils stochastiques pour déterminer l'allocation d'actifs a l'avantage de permettre de comparer des coûts d'option, il précise :

« Le résultat dépend totalement du jeu de *scenarii* utilisé, mais ça permet d'avoir un critère objectif : même si ça ne veut rien dire, c'est cohérent » (Entretien, Janvier 2016).

Les modèles permettent ainsi de justifier les pratiques de ceux qui s'y réfèrent (Boltanski et Thévenot, 1991) et, comme outil de justification, les grandeurs utilisées pour caractériser l'aléa sont d'autant plus efficaces qu'elles ne pourront jamais, rétrospectivement, être démenties. En situation d'aléa, une grandeur statistique n'est pas falsifiable (Frezal, 2015a). La modélisation d'une situation d'aléa, par définition, n'est en effet pas prédictive : en contexte d'incertitude, on ne peut entièrement exclure que certains scénarios défavorables se réalisent — et leur avènement n'est pas susceptible de remettre en cause la qualité de la décision prise *ex ante*. Celui qui utilise les modèles s'appuie sur un outil qui a formalisé l'événement adverse et ses conséquences éventuelles comme la conséquence d'une « faute à pas de chance ». À l'inverse, si l'on explicite, qualitativement, un ensemble de scénarios adverses afin de mener un raisonnement sur cette base, on s'expose au risque de la critique *ex post*, lorsque se produira un jeu de circonstances ne correspondant à aucun des scénarios de travail. Les statistiques font disparaître cette explicitation et permettent, dès lors, d'éviter la mise en critique et la nécessité de se justifier *ex post*. On le voit dans cette discussion informelle qui s'engage juste avant une réunion portant sur des choix d'allocation d'actifs. L'ancien CIO (*chief investment officer*) d'un grand groupe d'assurance européen, qui avait effectué toute sa carrière dans la gestion d'actifs, intervient pour expliquer qu'il préfère fonder sa décision sur l'optimisation de statistiques car « on a déjà essayé de raisonner sur des scénarios et on s'est toujours planté ! » (Journal de terrain, Janvier 2013). En fondant ses décisions sur trois ou quatre scénarios précis, le CIO constate son échec : aucun d'entre eux ne se réalise jamais. Il préfère donc se fonder sur la projection de mille scénarios stochastiques aléatoires, qui lui permettent de dégager une moyenne dont il pourra dire, *ex post*, qu'elle incluait le scénario qui est effectivement advenu. Comme le suggère Michael Power (2009), pour les

acteurs qui doivent prendre des décisions en situation d'incertitude, l'une des vertus des techniques quantitatives tient à ce qu'elles permettent de se soustraire, au moins partiellement, à la critique.

Dans certains cas, le recours aux grandeurs statistiques en situation d'aléa peut aller jusqu'à renvoyer à une forme de cynisme ou de manipulation délibérée.

Un membre d'une Direction des risques raconte ainsi :

« J'ai été interviewé par [un cabinet d'audit et de conseil] qui faisait une étude de marché sur les *best practices* d'allocation stratégique d'actifs. À la fin, je leur ai dit qu'au fond, plutôt que de fournir des indicateurs risque/espérance de rendement, on ferait mieux de fournir un rendement et la marge d'erreur associée. Ils m'ont alors regardé, mi-stupéfaits mi-inquiets, en s'exclamant : "mais si on fait cela, on va perdre toute crédibilité !" » (Entretien, Janvier 2016).

Les grandeurs statistiques, lorsqu'elles sont mises en regard, sont plus ou moins immédiatement convaincantes, semblent considérer les consultants : le couple risque/rendement ne soulève pas d'objection, alors que celui qui apparie une estimation déterministe et sa marge d'erreur verra sa pertinence opérationnelle trop aisément remise en cause. L'enjeu n'est plus ici d'y croire ou pas (les limites de l'outil sont connues, manifestement), mais de faire en sorte que l'interlocuteur puisse continuer d'y croire (« on va perdre toute *crédibilité* ! », nous soulignons).

4. Conclusion

Nous avons souhaité, dans cet article, contribuer à la connaissance d'un segment à cette heure encore peu exploré de la finance contemporaine : celui des assurances. Si l'on se fonde sur le traceur que constitue la circulation d'un raisonnement particulier, l'amalgame des situations d'aléa et des situations d'hétérogénéité, le secteur assurantiel, longtemps autonome, apparaît aussi comme subordonné : l'amalgame n'y pénètre que bien après qu'il s'est fait une place au cœur des pratiques des acteurs intervenant sur les marchés financiers

ou dans les activités bancaires, et c'est à partir des points de contact qu'ils entretiennent avec ces activités financières que les assureurs importent — en la radicalisant — cette forme de raisonnement. La radicalité de ce transfert renvoie sans aucun doute à la place qu'occupe l'assurance dans la circulation de cette institution : les modèles que vont mettre en place les assureurs bénéficient des sophistications préalablement apportées aux différents modèles, et reconduisent par ailleurs certains déplacements que d'autres avaient fait avant eux — comme le fait d'utiliser le calcul stochastique comme un instrument de gestion des risques. Mais elle tient également aux spécificités économiques de l'activité assurantielle, dont les contrats présentent des spécificités telles qu'ils offrent à l'amalgame un territoire plus étendu encore. C'est dire que le mimétisme qui est ici à l'œuvre n'a rien de mécanique, et que dans son transfert l'institution s'amende et se transforme : parce qu'elle n'existe que d'être mobilisée dans les pratiques de ceux qui s'en saisissent, elle dispose d'une plasticité bien supérieure à ce que certains schémas néo-institutionnalistes ont parfois pu présenter (Powell et Colyvas, 2008). En l'occurrence, les modèles qui s'appuient sur le calcul stochastique, dans leur version théorique initiale, ne mettent pas en jeu l'amalgame aléa/hétérogénéité. C'est parce que ces modèles sont mobilisés — de manière *a priori*, c'est-à-dire en étant soustraits aux jugements des faits — sur des territoires qui n'étaient initialement pas les leurs qu'ils en viennent à engager l'amalgame des situations d'aléa et d'hétérogénéité. L'étude de l'institutionnalisation de l'amalgame aléa/hétérogénéité permet à cet égard de faire retour sur certains débats classiques de la littérature néo-institutionnaliste. On voit, tout d'abord, si l'on tente de rendre compte du caractère partagé de l'institution, que les différents mécanismes qu'évoquaient jadis Paul DiMaggio et Walter Powell (1983) et qui continuent d'organiser cette littérature (voir par exemple Boxenbaum et Jonsson, 2008), jouent un rôle très inégal : le mimétisme et l'emprise normative ne parviennent, en l'occurrence, à expliquer l'appropriation de ce raisonnement que pour un nombre très limité d'acteurs (certes dominants). Ce n'est qu'en faisant intervenir des logiques

coercitives, avec la mise en place de Solvabilité II, que l'amalgame se diffuse à l'ensemble des acteurs — et l'on retrouve ici un résultat que C. Baud et E. Chiapello (2015) avaient mis en évidence pour le secteur bancaire : le rôle déterminant de la puissance publique dans la diffusion de certaines institutions, y compris des plus enclines à propager les logiques de la financiarisation.

Reconstituer la variété des pratiques qui font jouer à cet amalgame le rôle d'une institution permet, enfin, de revenir sur la conception de l'acteur qui sous-tend les travaux (néo-)institutionnalistes (Hwang et Powell, 2005). La littérature semble y être distribuée autour de deux pôles, très classiques et également insatisfaisants : celui qui voit dans l'acteur qui s'appuie sur l'institution un « idiot culturel » (DiMaggio et Powell, 1991), soumis comme un animal pavlovien aux injonctions unilatérales des institutions qui l'enserment, et celui qui au contraire en fait un manipulateur cynique pour qui le recours à l'institution n'est au fond rien d'autre que l'alibi d'un calcul. Les pratiques qui font jouer à l'amalgame aléa/hétérogénéité le rôle d'une institution parcourent, on l'a vu, tout le spectre qui s'étend entre ces deux pôles. Pour reconstituer les dynamiques institutionnelles, l'enjeu n'est donc pas d'assigner les observations à l'un ou l'autre de ces pôles, mais de parcourir l'espace qui s'établit entre eux en reconstituant la variété des motifs qui poussent à faire jouer un même rôle à un même raisonnement.

Références

- Autorité de Contrôle des Assurances et des Mutuelles (ACAM), 2008, Rapport d'activité 2008 – Les chiffres du marché français de l'assurance, 41 p.
- Baud, C., Chiapello, E., 2015, « Comment les firmes se financient. Sous le marché, les règles ? Le cas de la financiarisation du crédit bancaire », *Revue française de sociologie*, vol. 56, n° 3, p. 439-468.
- Beunza, D., Stark, D., 2004, « Tools of the trade. The socio-technology of arbitrage in a Wall-Street trading room », *Industrial and corporate change*, vol. 13, n° 2, p. 369-400.

- Boltanski, L., Thévenot, L., 1991, *De la justification*, Gallimard, Paris.
- Boxenbaum, E., Jonsson, S., 2008, « Isomorphism, diffusion and decoupling », *in* Greenwood, R., Oliver, C., Sahlin, K., Suddaby, R. (dir.), *The SAGE Handbook of Organizational Institutionalism*, Sage, Londres, p. 78-98.
- Breslau, D., Yonay, Y., 1999, « Beyond metaphor. Mathematical models in economics as empirical research », *Science in context*, vol. 12, n° 2, p. 317-332.
- Chauvin, P.-M., 2011, « Extension du domaine de la note. Robert Parker comme saillance du marché des grands vins de Bordeaux », *in* François, P. (dir.), *Vie et mort des institutions marchandes*, Les Presses de Sciences Po, Paris, p. 79-107.
- Chiapello, E., 2005, « Les normes comptables comme institutions du capitalisme. Une analyse du passage aux IFRS en Europe à partir de 2005 », *Sociologie du travail*, vol. 47, n° 3, p. 362-382.
- Chiapello, E., Gilbert, P., 2009, « La gestion comme technologie économique », *in* Steiner, P., Vatin, F. (dir.), *Traité de sociologie économique*, Presses universitaires de France, Paris.
- Davis, G. F., Diekmann, K. A., Tinsley, C. H., 1994, « The decline and fall of the conglomerate firm in the 1980s?: the deinstitutionalization of an organizational form », *American sociological review*, vol. 59, n° 4, p. 547-570.
- Diaz, F., 2005, « L'observation participante comme outil de compréhension du champ de la sécurité », *Champ pénal/ Penal field*, vol. II, en ligne : <http://journals.openedition.org/champpenal/79>.
- DiMaggio, P. J., 1988, « Interest and agency in institutional theory », *in* Zucker, L. G. (dir.), *Institutional patterns and organizations*, Ballinger, Cambridge, p. 3-32.
- DiMaggio, P. J., Powell, W. W., 1983, « The iron cage revisited: institutional isomorphism and collective rationality in organizational fields », *American sociological review*, vol. 48, n° 2, p. 147-160.

- DiMaggio, P. J., Powell, W. W., 1991, « Introduction », in DiMaggio, P. J., Powell, W. W. (dir.), *The new institutionalism in organizational analysis*, University of Chicago Press, Chicago, p. 1-38.
- Dubois, S., François, P., 2013, « Seeing the world through common lenses? The case of french contemporary poetry », in Beckert, J. et Musselin, C. (dir.), *Constructing quality. The classification of goods in the economy*, Oxford, Oxford University Press, p. 174-193.
- Fourcade, M., Khurana, R., 2013, « From social control to financial economics. The linked ecologies of economics and business in twentieth century America », *Theory and society*, vol. 42, n° 2, p. 121-159.
- François, P., 2011, « Puissance et genèse des institutions : un cadre analytique », in François, P. (dir.), *Vie et mort des institutions marchandes*, Les Presses de Sciences Po, Paris, p. 39-79.
- François, P., 2015, « La fabrique sociale d'une *tabula rasa* : le lancement de *Solvency II* », Working paper PARI, 59 p., <http://www.chaire-pari.fr/wp-content/uploads/2015/11/Francois-S2-Tabula-rasa.pdf>
- François, P., Frezal, S., 2016, « Pourquoi utilisons-nous des modèles “faux” ? », Working paper PARI, 23 p., en ligne : <http://www.chaire-pari.fr/wp-content/uploads/2016/09/Genese-alea-heterogeneite-9-Septembre-2016-MEP.pdf>
- François, P., Lemerrier, C., 2016, « Une financiarisation à la française (1979-2009). Mutations des grandes entreprises et conversion des élites », *Revue française de sociologie*, vol. 57, n° 2, p. 269-320.
- Frezal, S., 2015a, « Aléa et hétérogénéité : l'amalgame tyrannique », Working paper PARI, 43 p., en ligne : http://www.chaire-pari.fr/wp-content/uploads/2015/11/concepts_amalgame-tyrannique-25112015.pdf
- Frezal, S., 2015b, « Le couple risque/rendement : une chimère ? », Working paper PARI, 30 p., en ligne : http://www.chaire-pari.fr/wp-content/uploads/2015/11/Le-couple-risque_rendement-Une-chimere-19102015.pdf

- Frezal, S., 2016, « De quoi Solvabilité 2 est-il le nom ? », *Working paper PARI*, 54 p., en ligne : http://www.chaire-pari.fr/wp-content/uploads/2016/03/De-quoi-Solvabilité-2-est-elle-le-nom_mars-2016.pdf
- Garud, R., Jain, S., Kumaraswamy, A., 2002, « Institutional entrepreneurship in the sponsorship of common technological standards: the case of Sun Microsystems and Java », *Academy of Management Journal*, vol. 45, n° 1, p. 196-214.
- Hubert, H., Mauss, M., 1985, « Esquisse d'une théorie générale de la magie », in Mauss, M., *Sociologie et anthropologie*, Presses universitaires de France, Paris, p. 3-142.
- Hwang, H., Powell, W. W., 2005, « Institutions and entrepreneurship », in Alvarez, S., Agarwal, R. et Sorenson, O. (dir.), *Handbook of entrepreneurship research*, Springer, Kluwer, p. 179-210.
- Jarzabkowski, P., Bednarek, R., Spee, P., 2015, *Making a market for acts of god*, Oxford University Press, Oxford.
- Jovanovic, F., 2012, « Finance in modern economic thought », in Knorr-Cetina, K., Preda, A. (dir.), *The Oxford handbook of the sociology of finance*, Oxford University Press, Oxford.
- Latour, B. et Woolgar, S., 1988, *La vie de laboratoire*, La Découverte, Paris.
- Lawrence, T. B., Suddaby, R., 2006, « Institutions and institutional work », in Clegg, S. R., Hardy, C., Lawrence, T. B., Nord, W. R. (dir.), *Handbook of organization studies*, Sage, Londres, p. 215-254.
- Linsmeier, T., Pearson, N., 2000, « Value at Risk », *Financial analysts Journal*, vol. 56, n° 2, p. 46-67.
- Lizé, W., 2009, « Entretiens, directivité et imposition de problématique. Une enquête sur le goût musical », *Genèses*, n° 76, p. 99-115.
- Lounsbury, M., Hirsch, P. M. (dir.), 2010, *Markets on trial*, Emerald Group Publishing, Londres.

- Markowitz, H., 1952, « Portfolio selection », *The Journal of Finance*, vol. 7, n° 1, p. 77-91.
- MacKenzie, D., 2003, « An equation and its worlds. Bricolage, exemplars, disunity and performativity in financial economics », *Social studies of science*, vol. 33, n° 6, p. 831-868.
- MacKenzie, D., 2006, *An engine, not a camera*, MIT Press, Cambridge.
- MacKenzie, D., Millo, Y., 2003, « Constructing a market: the historical sociology of a financial derivatives exchange », *American Journal of Sociology*, vol. 109, n° 1, p. 107-145.
- MacKenzie, D., Spears, T., 2014, « “The formula that killed Wall Street”: The Gaussian Copula and modelling practices in investment banking », *Social studies of sciences*, vol. 44, n° 3, p. 393-417.
- Merlus, S., Pecqueux, O., 2000, *Les garanties plancher des contrats d'assurance-vie en UC : tarification et couverture*, Mémoire d'actuariat de l'ENSAE.
- Oliver, C., 1992, « The antecedents of deinstitutionalization », *Organization Studies*, vol. 13, p. 563-588.
- Ollivier, C., 2011, « Naissance et survie d'une institution. La qualification professionnelle des architectes d'intérieur », in François, P. (dir.), *Vie et mort des institutions marchandes*, Les Presses de Sciences Po, Paris, p. 195-223.
- Peneff, J., 1996, « Les débuts de l'observation participante ou les premiers sociologues en usine », *Sociologie du travail*, vol. 38, n° 1, p. 25-44.
- Powell, W. W., Colyvas, J. A., 2008, « Microfoundations of institutional theory », in Greenwood, R., Oliver, C., Suddaby, R., Sahlin-Anderson, K. (dir.), *The Sage handbook of organizational institutionalism*, Sage, Thousand Oaks.
- Power, M., 2009, « The risk management of nothing », *Accounting, Organizations and Society*, vol. 34, n° 6-7, p. 849-855.
- Rao, H., Sivakumar, K., 1999, « Institutionnal sources of boundary-spanning structures: the establishment of investor relations departments in the Fortune 500 industrials », *Organization Science*, vol. 10, n° 1, p. 27-42.

- Rao, H., Monin, P., Durand, R., 2003, « Institutional Change in Toque Ville: Nouvelle Cuisine as an Identity Movement in French Gastronomy », *American Journal of Sociology*, vol. 108, n° 4, p. 795-843.
- Rona-Tas, A., 2010, « The role of ratings in the subprime mortgage crisis », in Lounsbury, M., Hirsch, P. M. (dir.), *Markets on trial*, Emerald Group Publishing, Londres.
- Rubinstein, M., 2002, « Markowitz's portfolio selection. A fifty-year retrospective », *The journal of finance*, vol. 62, n° 3, p. 1041-1045.
- Shafer, G., Vovk, V., 2001, *Probability and Finance: It's only a game*, Wiley, New York.
- Underhill, G. R. D., Zhang, X., 2008, « Setting the rules. Private power, political underpinnings, and legitimacy in global monetary and financial governance », *International Affairs*, vol. 84, n° 3, p. 535-554.
- Zelizer, V., 1979, *Morals and markets*, Columbia University Press, New York.

Chapitre 6 : Le basculement vers une logique marchande

De quoi Solvabilité 2 est-il le nom ?

Introduction

Solvabilité 2 : quel objectif ?

La directive dite « Solvabilité 2 » (*European Parliament and Council, 2009*) a été développée en Europe entre la fin des années 90 et le milieu des années 2010, entrant en vigueur début 2016. Faisant table rase (François, 2015) du système prudentiel précédent, dit Solvabilité 1, elle a défini un nouveau cadre réglementaire encadrant l'exercice de l'activité d'assurance.

Les enjeux liés à la régulation d'un secteur sont divers et concernent plusieurs types d'acteurs : les assurés et le régulateur qui a pour mission de les protéger, que nous assimilerons ici⁷⁷ ; les organismes régulés et leurs contraintes en termes de réputation, de rentabilité, etc. ; l'Etat et ses objectifs macroéconomiques ; les investisseurs et notamment les actionnaires ; les organismes privés « facilitateurs de confiance », tels que les commissaires aux comptes ou les agences de notation ; les fournisseurs de services dont l'activité dépend de la réglementation et de son évolution, tels que les cabinets de conseil et les sociétés de services informatiques, etc. Chacune de ces catégories de parties prenantes est susceptible d'être affectée par le design d'une régulation. En fonction de l'objectif de celle-ci, les enjeux de l'une ou l'autre de ces parties-prenantes seront censés être centraux, périphériques, ou bien sans objet. Par exemple, la régulation en matière de dimensionnement des équipements de sûreté nucléaire est censée avoir pour objectif central la fiabilité, avoir un enjeu annexe de coût et être indépendante des enjeux de lisibilité des comptes d'EDF pour les investisseurs ; réciproquement, la régulation en matière de qualité de l'information fournie aux investisseurs, même si elle s'applique à Danone, n'a pas pour objectif de contribuer à la qualité sanitaire de leurs yaourts.

⁷⁷ Nous ne nous plaçons en effet pas ici dans l'économie politique des éventuels conflits d'intérêts, individuels ou collectifs, des agents de l'autorité de régulation, mais dans un cadre où les objectifs du régulateur sont assimilables aux objectifs de la régulation prudentielle, c'est-à-dire la préservation des intérêts de la collectivité des assurés.

Dans le domaine de la régulation du secteur financier et notamment des assurances, l'objectif d'une régulation prudentielle est *a priori* de protéger les clients des organismes financiers contre le risque de défaut de ceux-ci en encadrant leur prise de risque. De fait, l'existence d'un contrôle prudentiel résulte de la combinaison d'une inversion du cycle de production (les assureurs encaissent la prime versée par l'assuré, puis devront ultérieurement fournir le service, c'est-à-dire tenir leur promesse de remboursement) et d'une forte asymétrie d'information quant à la santé financière et la qualité de gestion du vendeur, combinaison spécifique à ce secteur qui rend nécessaire une intervention tierce ayant pour finalité de permettre un certain niveau de confiance du consommateur envers l'organisme financier. Ainsi, Solvabilité 1 avait pour objectif de garantir la « protection des assurés » dans le marché commun européen : si une réglementation *prudentielle* peut avoir des impacts sur l'ensemble des acteurs que nous avons évoqués, sa finalité est *a priori* la protection des intérêts des assurés, les intérêts des autres parties prenantes relevant, ou non, d'autres textes réglementaires.

La directive Solvabilité 2 marque toutefois une évolution par rapport à ce cadre, introduisant aux côtés de la finalité historique de la régulation prudentielle deux autres missions, « *la stabilité financière et la stabilité et l'équité des marchés* », en les lui subordonnant explicitement :

(16) Le principal objectif de la réglementation et du contrôle en matière d'assurance et de réassurance est de garantir la protection adéquate des preneurs et des bénéficiaires. [...] La stabilité financière et la stabilité et l'équité des marchés constituent d'autres objectifs de la réglementation et du contrôle en matière d'assurance et de réassurance qui devraient être également pris en compte, *sans détourner cependant du principal objectif*. (Parlement européen et Conseil, 2009 ; je souligne)

Ce choix de coexistence d'objectifs peut s'expliquer par le fait que la protection des assurés repose, pour sa dimension quantitative, sur l'analyse financière des entreprises ; par conséquent, les deux objectifs de pertinence de l'information fournie aux marchés et de protection des assurés pourraient reposer sur des

grilles de lecture partagées, sur des moyens mutualisables : la réglementation prudentielle serait susceptible de bénéficier non seulement aux assurés, mais également aux investisseurs finançant l'activité d'assurance.

Mais peut-on courir plusieurs lièvres à la fois ? Solvabilité 2 introduit des évolutions (suppression de la prudence dans l'évaluation des postes du bilan, valorisation des perspectives de profits futurs permettant de libérer du capital, etc.) dont on pourrait spontanément penser qu'elles ne s'inscrivent pas dans la poursuite de l'objectif réputé principal de protection des assurés *et iraient à son contre*. Il est donc nécessaire de décortiquer plus finement les choix retenus afin de comprendre si la hiérarchie des objectifs que s'était fixé Solvabilité 2 a été respectée. Quel objectif a eu la prééminence lors de la conception du système et lequel a été, consciemment ou non, la variable d'ajustement dans la transition entre les deux systèmes prudentiels ? Lequel a bénéficié d'une amélioration lors du passage de Solvabilité 1 à Solvabilité 2 et lequel est susceptible d'avoir subi une dégradation ?

Cette interrogation sur les objectifs peut être formulée, de façon duale, sur les acteurs ayant bénéficié de cette réforme. De fait, on ne peut faire l'économie d'un questionnement sur le contexte dans lequel s'inscrit l'évolution de la réglementation : pourquoi a-t-on considéré que l'ancienne réglementation n'était plus adaptée ? Les choix de valorisation des éléments du bilan prudentiel sous Solvabilité 2 sont indissociables des réflexions et évolutions comptables parallèles. Le responsable de l'unité assurance à la Commission européenne sous la mandature de qui ces principes ont été actés était auparavant responsable de l'unité en charge des règles comptables qui a porté les normes IFRS généralisant en Europe l'usage de la « juste valeur » (François, 2015). Si l'industrie de l'assurance était alors restée en dehors du champ d'application de ces normes, Solvabilité 2 les y fait entrer de plain-pied *via* la comptabilité

prudentielle.⁷⁸ La mise en place de ce système pourrait donc être lue comme un mécanisme ayant conduit à contourner l'exclusion du secteur de l'assurance de la nouvelle vision comptable portée par les normes IFRS et le faisant basculer dans cette nouvelle logique où « *la vision de l'organisation n'est plus tant celle d'une entreprise produisant une marchandise que celle d'une entreprise étant elle-même une marchandise pouvant être vendue et achetée par les investisseurs* » (Chiapello 2014, je traduis). Ces acteurs n'étaient *a priori* pas ciblés pour être les premiers bénéficiaires d'une réforme de la régulation prudentielle. Dès lors, notre question émerge : sont-ils des bénéficiaires collatéraux, ou bien leurs enjeux ont-ils au contraire été placés au cœur de la réforme ? Dans la conception de son pilier quantitatif, le choix de ses principes et de ses métriques, dans ses conséquences opérationnelles sur le pilotage des compagnies, Solvabilité 2 représente-t-elle une régulation prudentielle améliorant la protection des assurés, ou bien une régulation de marché à des fins d'analyse financière ?

La trame et l'intuition

La première partie sera strictement descriptive : elle exposera les principes comptables, les principaux postes bilanciaux et l'architecture des contraintes quantitatives des deux systèmes prudentiels : Solvabilité 1 dans un premier temps, puis Solvabilité 2.

En nous appuyant sur cette description, il apparaîtra que Solvabilité 1 répond à *une cohérence d'objectif*. S'il s'agit d'une sédimentation de différentes composantes, de différentes strates de mesures et de protections, cet ensemble est ancré dans une vision industrielle, c'est-à-dire contingente au produit vendu. Les appréciations qui en résultent sont cohérentes les unes avec les autres en ce qu'elles partagent la même finalité, centrée sur la garantie de l'engagement pris envers l'assuré.

⁷⁸ De fait, le système quantitatif déployé dans Solvabilité 2 s'appuie sur, et justifie l'utilisation de, la *juste valeur* au sein d'un ensemble {valorisation du bilan ; détermination de l'exigence de capital} cohérent.

Dans la seconde partie, nous décortiquerons l'interprétation qui peut être faite des métriques retenues dans Solvabilité 2. Ceci permettra de mettre en évidence que Solvabilité 2 correspond davantage à *un système de pensée* : son ambition de cohérence est celle d'une grille d'analyse se voulant neutre, comme dénuée de finalité. Il apparaîtra qu'il s'agit d'un cadre théorique ancré dans une vision financière, centré sur le pilotage de la valeur pour l'actionnaire.

La troisième partie décrira comment cette grille d'analyse pourrait faire évoluer la représentation qu'on se fait d'une société d'assurance, l'état d'esprit des régulateurs, ainsi que les modes de pensée et les pratiques de gestion et des organismes d'assurance.

La lecture de la première partie n'est pas nécessaire au lecteur familier de la régulation prudentielle des assurances en Europe qui connaît déjà les systèmes dits « Solvabilité 1 » et « Solvabilité 2 ».

Caveat

Les exigences quantitatives ne constituent pas l'unique aspect d'une régulation prudentielle. Celle-ci passe bien évidemment par le droit du contrat (prévenir les organismes de se livrer à une surenchère concurrentielle de promesses intenables) et, en termes de droit des entreprises, par des contraintes autres que quantitatives.

Ainsi, avant Solvabilité 2, les entreprises devaient chaque année mener une réflexion sur les risques qu'elles prenaient et leur capacité à maintenir leur solvabilité, et formaliser celle-ci à destination de l'autorité de contrôle par la rédaction d'un « rapport de solvabilité » et d'autres documents, tels qu'un « rapport de contrôle interne » ou encore un « rapport sur la politique de placement »⁷⁹. Parallèlement, elles devaient envoyer des données aux autorités de contrôle *via* la production de reportings (états T pour les reportings trimestriels et états C pour les reportings annuels)⁸⁰.

⁷⁹ Code des Assurances, L. 322-2-4, R. 336-1 et L.143-6. Ce dernier rapport est communiqué à tout assuré qui le demande.

⁸⁰ Code des Assurances A. 344-13 et A.344-10.

Solvabilité 2 formalise ces deux types d'outils en créant le « pilier 2 » et le « pilier 3 ». Le pilier 2 rassemble les exigences dites qualitatives : d'une part, comme le système précédent, elle demande aux organismes d'assurance d'évaluer leurs risques (*via* la production d'un rapport dit ORSA, *Own Risk and Solvency Assessment*, équivalent du rapport de solvabilité dans le système précédent) et, d'autre part, elle leur enjoint en outre de formaliser leur gouvernance, incitant à la formalisation de process de décision, la production de minutes des comités, etc., et imposant notamment la formalisation de certaines fonctions dites *fonctions clés* (contrôle interne, conformité, etc.). Le pilier 3 rassemble deux jeux d'exigences de reporting, d'une part à destination de l'autorité de contrôle comme dans le système précédent et, d'autre part, à destination du public (analystes, pairs, agences de notation, assurés, etc.). Cette deuxième composante, est assez largement nouvelle⁸¹. Elle s'inspire de la régulation bancaire et sa promotion repose sur l'idée que, en s'offrant à l'analyse du public, chaque organisme d'assurance, se sachant observé par les marchés, sera incité à être prudent pour inspirer confiance.

L'objet de cet article n'est pas de comparer les piliers 2 et 3 à leurs équivalents antérieurs : nous nous concentrerons ici sur les exigences quantitatives, qui constituent le « pilier 1 ».

1. Les faits : description des systèmes prudentiels

1.1 La pierre angulaire : le bilan

L'activité d'assurance consiste à vendre des promesses contingentes à un état futur du monde. Par exemple, la promesse du versement d'une indemnisation contingente à la réalisation d'un sinistre (assurance non vie), ou la promesse du versement d'un flux de revenus contingente à la survie de l'assuré (assurance retraite). La réglementation prudentielle est une réponse à l'asymétrie d'information entre l'assuré et l'assureur quant à la capacité de ce dernier à

⁸¹ Même si, juridiquement, chacun pouvait déjà obtenir auparavant, sur simple demande et pour le prix d'un timbre, la liste des placements d'un assureur, avec Solvabilité 2, la quantité de données ouvertes au public et les modalités de transmission d'information changent de dimension.

tenir ses promesses : elle a pour objectif de favoriser la confiance envers l'assureur, condition nécessaire d'existence d'un marché des produits d'assurance. Sa finalité est de fournir à l'assuré une garantie suffisante quant à la santé financière de l'assureur.

Elle repose donc sur une évaluation de ses engagements, des actifs qu'il détient pour y faire face, et de l'écart entre les deux qui représente un matelas de sécurité permettant d'absorber la survenance d'événements adverses tels qu'une chute de la valeur de ses actifs ou une réévaluation de l'estimation de ses engagements. Le bilan d'une société d'assurance est ainsi constitué essentiellement de trois composantes :

- les provisions (*reserves*), qui correspondent à l'évaluation des engagements envers les assurés, et constituent l'essentiel de son passif,
- les investissements, qui représentent l'essentiel de l'actif et sont constitués majoritairement d'obligations ainsi que, dans une moindre mesure, d'actions, d'immobilier et de trésorerie,
- les fonds propres qui correspondent, grossièrement, à l'écart entre les deux.

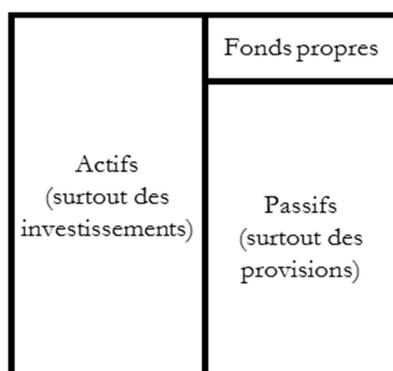


Schéma 1 : bilan schématique d'un organisme d'assurance⁸²

Solvabilité 1 et Solvabilité 2 ont adopté des approches d'évaluation de chacun de ces blocs radicalement différentes.

⁸² En Solvabilité 1, les fonds propres représentent de l'ordre de 4% du bilan des assureurs vie et de l'ordre de 20% du bilan des assureurs non vie (ACPR, 2014b, p83).

1.2 Une mise en regard usuelle devant être approfondie

Le contraste entre Solvabilité 1 et Solvabilité 2 est double : il porte d'une part sur ce bilan, donc l'évaluation des engagements, des actifs et des fonds propres disponibles, et, d'autre part, sur la détermination de l'exigence de capital, c'est-à-dire sur le montant minimal de fonds propres requis.

Du point de vue bilanciel, le discours usuel oppose un bilan Solvabilité 1 « comptable », terme chargé d'opprobre, auquel il est reproché d'être conventionnel et partant non objectif voire erroné⁸³, à un bilan Solvabilité 2 dit « économique », en valeur de marché, réputé représenter une vision juste, c'est-à-dire intrinsèquement pertinente, objective, non biaisée.

Du point de vue de l'exigence de capital, le discours commun oppose les traditionnelles exigences de Solvabilité 1, considérées comme non *risk-based*, « arbitraires » car ne découlant pas d'un calcul, à une vision censée être fondée sur le risque⁸⁴, donc par principe préférable tant en termes d'efficacité économique (les assurés bénéficient tous de la même protection, les compagnies plus risquées devant immobiliser davantage de fonds propres) que de morale sociale (celui qui est joueur doit payer en immobilisant davantage de capital).

Ces deux évolutions sont synthétisées dans un des tout premiers paragraphes de la directive, qui indique :

(15) Conformément aux derniers développements des travaux de l'Association internationale des contrôleurs d'assurance, du Conseil des normes comptables internationales et de l'Association actuarielle internationale concernant la gestion des risques, ainsi qu'à l'évolution récente observée dans les autres

⁸³ « Non objectif » car d'autres conventions auraient pu être prises, « erroné » par exemple parce que, lors de la comptabilisation d'un actif en coût historique, (i) cette valeur figée reflète mal les variations de la richesse détenue par l'organisme au fil du temps, et (ii) deux titres identiques peuvent, s'ils ont été achetés à des dates différentes, être valorisés différemment.

⁸⁴ Lorsque le terme « risque » est ainsi utilisé dans le cadre de l'exigence de capital, cela fait référence à l'aléa auquel l'organisme est soumis et non aux risques des clients que l'assureur mutualise et qui sont statistiquement provisionnés.

secteurs financiers, *il convient d'adopter une approche économique fondée sur le risque [...]*.(EPC, 2009, je souligne)

Le caractère intrinsèquement positif de cette approche est internalisé par les acteurs promouvant Solvabilité 2. Ainsi, D. Kessler, P-DG de la Scor, affirme que :

« Solvabilité 2 permet l'introduction de cette dimension économique et basée sur les risques. Pour un assureur, on ne peut pas imaginer un système plus intelligent ». (14 octobre 2014, Journée *Scor Global Life – Solvabilité 2*)

Toutefois, ce discours présente des fissures. Premièrement, si la valorisation des actifs en valeur de marché, quoique longuement débattue au sein des autorités comptables en général (Chiapello 2005), est parvenue à se faire naturaliser comme étant une valorisation « juste », désormais réputée naturelle par opposition à conventionnelle, il ne semble pas possible de considérer la valorisation d'un passif d'assurance en valeur dite « de marché » comme ne résultant pas de conventions. A titre d'exemple, la description des méthodes de valorisation des passifs en valeur de marché nécessite 70 pages dans les spécifications techniques du QIS 5 (*European Commission*, 2010), spécifications qui laissent pourtant encore ouvertes de nombreuses interprétations et d'innombrables hypothèses (cf. *infra*).

Deuxièmement, avant même l'entrée en vigueur de la directive, un bilan « économique » n'apparaissait déjà plus comme reflétant « la vérité ». Ainsi, en 2014, dans le cadre de la réflexion sur le régime prudentiel auquel devraient être soumis les IRP, le concept de « *holistic balance sheet* » a été inventé. Ce dernier prend en compte des éléments qui ne l'étaient pas dans le bilan « économique » défini par Solvabilité 2 pour les assureurs classiques :

« The holistic balances sheet captures these national specificities by allowing IORPs to include explicitly all security and benefit adjustment mechanisms » (EIOPA, 2014)

montrant en négatif si besoin était que le bilan dit « économique » n'offre qu'une vision incomplète de l'entreprise et est un choix de représentation parmi d'autres.

Troisièmement, Frezal (2016) indique que les mesures de risque sur lesquelles repose l'exigence de fonds propres sont trop instables et imprécises pour pouvoir être considérées comme des indicateurs de risque (fussent-ils relatifs) opérationnellement fiables.

Dans ce cadre, commençons par préciser la portée et les limites du discours usuel en présentant plus finement chacun de ces deux systèmes.

1.3 Solvabilité 1

Solvabilité 1 repose sur la comptabilité sociale et un principe de prudence⁸⁵. En non-vie, la façon dont celui-ci s'incarne par un jeu de provisions *ad hoc* est précisée au niveau national. Nous prenons dans la suite l'exemple de la France, l'état d'esprit et les principaux postes étant partagés par les différents pays. En vie, les textes européens sont davantage précis mais laissent toujours au niveau national le soin de préciser.

*Provisionnement non vie*⁸⁶

Le principal poste au bilan d'une société d'assurance non vie (c'est-à-dire proposant des couvertures non liées à la durée de la vie humaine : auto, multirisque habitation, etc.) est constitué par les *provisions pour sinistres à payer (PSAP)*. Ce poste fait référence aux sinistres ayant déjà eu lieu, mais que l'assureur n'a pas encore indemnisés (ou du moins pas totalement). Il s'agira par exemple du dégât des eaux survenu fin décembre qui ne lui aura pas encore été déclaré, de l'incendie survenu en octobre pour lequel les rapports d'expert ne sont pas encore obtenus, ou bien du piéton renversé il y a trois ans couvert en responsabilité civile et dont l'état de santé non encore stabilisé ne permet pas au corps médical et au juge de fixer le montant définitif de l'indemnité. Pour les dossiers connus, l'assureur peut estimer une valeur de l'indemnité qui sera versée mais il n'a pas encore tous les éléments en main, et certains dossiers

⁸⁵ En vie, directive 2002/83 – art. 20 (EPC 2002); en non-vie, directives 73/239 – art. 15 puis 92/49 – art. 17 (EC 1973 et EC 1992).

⁸⁶ Code des Assurances, R. 331-6.

ne sont pas encore connus. Ainsi, ce poste majeur est une évaluation statistique.

L'autre poste significatif correspond aux *provisions pour primes non acquises (PPNA)*. Ce poste correspond au fait que, si un contrat a été signé le premier octobre par exemple, les trois quarts de la prime encaissée correspondent à une garantie fournie pour l'année à venir : ils ne peuvent être placés en résultat et doivent être provisionnés pour couvrir le montant qui sera statistiquement dépensé l'année suivante.

Le principe d'évaluation de ces provisions est un principe de « prudence », les provisions devant être « suffisantes », c'est-à-dire suffisantes pour permettre l'indemnisation des assurés (étant implicitement entendu que tel doit être le cas « même en cas de mauvaise surprise »).

- Concernant les PSAP, entièrement statistiques, l'appréciation des niveaux de prudence requis et existant est laissée à la responsabilité des dirigeants.
- Concernant les PPNA, la prudence est imposée. Ainsi, le montant de la provision est celui de la prime *pro-rata temporis* : comme une entreprise fait structurellement des bénéfices, cela signifie qu'on a tendance à surestimer le montant de la provision : on est ainsi prudent. Si tel n'est pas le cas, c'est-à-dire si les primes se sont révélées insuffisantes dans le passé ou bien si on anticipe une dégradation de la sinistralité dans le futur, alors d'autres provisions devront venir s'ajouter aux PPNA pour les compléter, comme des matelas de protection (e.g. *provision pour risque en cours, provision pour risques croissants*), toujours de façon *ad hoc*, par catégories de garanties.

*Provisionnement vie*⁸⁷

La quasi-totalité du passif d'une société d'assurance vie (c'est-à-dire commercialisant, essentiellement, des garanties d'épargne et de retraite), est

⁸⁷ Code des Assurances, R. 331-3.

constitué par les *provisions mathématiques* (PM). Elles correspondent au montant des versements déjà effectués et capitalisés ou, pour une rente en cours de versement, au montant actualisé qui sera statistiquement versé, calculé d'après des tables de mortalité prudentes et un taux d'actualisation prudent.

Ici encore, une prudence est explicitement requise⁸⁸ et, ici encore, des couches de prudence devront être ajoutées en constituant des provisions supplémentaires si nécessaire, notamment en cas de crainte que les rendements des actifs soient dans le futur trop faibles pour servir les taux de revalorisation garantis par l'assureur (par exemple la *provision pour aléas financiers* ou la *provision pour risque de taux*), ou encore pour s'assurer que l'assureur aura les ressources suffisantes pour non seulement servir les taux garantis mais également assurer la gestion des contrats (par exemple la *provision pour frais d'acquisitions reportés* ou la *provision globale de gestion*).

Surtout, la réglementation prudentielle est ici indissociable du droit du contrat. En effet, le meilleur moyen de tenir ses engagements est de s'engager peu. Ainsi, la réglementation interdit aux assureurs de proposer des taux garantis trop élevés. Le corollaire est qu'ils auront tendance à réaliser souvent des bénéfices significatifs : la réglementation les oblige donc parallèlement à reverser une part minimale (85 à 90%) de leurs bénéfices aux assurés. Ils disposent de huit années maximum pour cela et les montants sont durant cette période stockés dans la provision pour *participation aux excédents* (PPE).

*Actifs*⁸⁹

Les actifs correspondent essentiellement aux investissements réalisés à l'aide des primes collectées, et alors que les indemnités et les versements n'ont pas encore été réalisés. Là encore, un principe de prudence est à l'œuvre : on

⁸⁸ « Une évaluation prudente ne signifie pas une évaluation sur la base des hypothèses considérées les plus probables, mais doit tenir compte d'une marge raisonnable pour variations défavorables des différents facteurs en jeu. » dir. 2002/83, art. 20. 1. A. iii.

⁸⁹ Comptabilité sociale en France : Code des Assurances, R332-2, 3 et 19 à 21 et R. 331-3. Règles de dispersion au niveau européen : directives 92/49 – art. 20 et 22 (non vie) et 2002/83 – art. 22 et 24 (vie) (EC, 1992 et EPC, 2002).

retient donc le coût historique, donc sans prise en compte des plus ou moins-values latentes mais, afin de ne pas surestimer les actifs, des correctifs pourront être appliqués en cas de moins-values latentes. Il s'agit par exemple de la *provision pour dépréciation durable*, passée si un titre semble significativement dévalorisé, à l'instar de ce qu'on observe en comptabilité sociale générale, ou encore de la *provision pour risque d'exigibilité*, dans le cas où les actions et l'immobilier⁹⁰ pris dans leur ensemble sont en moins-value.

Au-delà des règles d'évaluation des actifs, des contraintes sont imposées pour éviter qu'une entreprise ne concentre trop ses placements. Ainsi des plafonds proportionnels à la taille des passifs sont fixés par catégorie d'émetteur (e.g. l'ensemble des actions ou de l'immobilier ne doit pas dépasser 65 ou 40% des provisions), ou par émetteur (l'ensemble des actions et obligations émises par telle société ne peut dépasser, par exemple, 5% des provisions)

Fonds propres disponibles⁹¹

Le montant de provision est ce qui détermine, par différence avec le total des actifs, le niveau de fonds propres de l'entreprise c'est-à-dire, lorsque le seuil « exigence minimale » est franchi, des mesures telles que le retrait d'agrément et, lorsque le seuil « 0 » est franchi, la mise en faillite.

Les fonds propres d'une société d'assurance correspondent, par construction, à l'écart entre les actifs et les passifs remboursables. Ils correspondent au capital social initialement réuni par les actionnaires auquel se sont ajoutés au fil des années les résultats obtenus nets des versements de dividendes⁹².

⁹⁰ On considère qu'une moins-value obligatoire est virtuelle puisque l'obligation est réputée pouvoir être détenue jusqu'à maturité.

⁹¹ Au niveau européen : directives 92/49 – art. 24 (non vie) et 2002/83 – art. 27 (vie).

⁹² Ainsi que diverses composantes, principalement, le cas échéant, les emprunts subordonnés (et, sur autorisation de l'autorité de contrôle, les plus-values latentes ; nous reviendrons sur ce point).

Exigences de capital⁹³

Les fonds propres doivent permettre d'absorber des résultats négatifs sans que ceux-ci aient pour conséquence directe et immédiate une insuffisance des actifs par rapport aux engagements, c'est-à-dire une faillite à l'occasion de laquelle les garanties envers les assurés ne seraient pas honorées. Leur montant minimum est déterminé de façon rustique, fonction dans une certaine mesure de l'activité de l'entreprise et de sa taille.

En non vie, deux références sont retenues : le chiffre d'affaires récent et le montant d'indemnités récent. A chacune, on applique un pourcentage forfaitaire contenant deux types de correctifs :

- afin de prendre en compte la plus grande stabilité des résultats liés à un portefeuille large où la loi des grands nombres est plus fiable, ce pourcentage est diminué au-delà d'un certain seuil,
- afin de prendre en compte la plus grande volatilité de certains types de garanties, ce pourcentage est augmenté (multiplié par 1,5) pour elles.

Enfin, on retient le plus élevé des deux montants, corrigé d'un paramètre reflétant, là encore de façon grossière, le fait que les programmes de réassurance mis en œuvre par l'assureur contribuent à l'immuniser plus ou moins contre les mauvaises nouvelles, à partir des observations passées et de façon fortement limitée, toujours dans un souci de prudence.

En vie, où les dangers auxquels est exposé un assureur relèvent davantage du stock d'engagements accumulés que du flux annuel de nouveaux engagements, les exigences de capital sont principalement fondées sur le montant des provisions mathématiques. Là encore, le caractère plus ou moins risqué des engagements pris par l'assureur est reflété par des calibrages forfaitaires : en France par exemple, il est quatre fois plus élevé lorsque l'assureur garantit le montant investi que lorsque tel n'est pas le cas (c'est-à-dire pour les contrats en unité de compte). Enfin, si l'assureur garantit un montant spécifique en cas de

⁹³ Au niveau européen : directives 73/239 - art. 16 (non vie) et 2002/83 – art. 28 (vie).

décès, l'exigence de capital est également accrue d'une certaine proportion de ces montants cumulés.

Synthèse

L'architecture prudentielle de Solvabilité 1 repose donc sur trois pans :

- des contraintes explicites, limitant le champ d'action des assureurs (contraintes de dispersion à l'actif, limitation des garanties proposées au passif),
- des règles d'évaluation du bilan, dont le principe de calcul est qu'elles doivent permettre l'indemnisation des assurés, en tendant à surestimer le montant des provisions et sous-estimer le montant des actifs,
- des règles de calcul de l'exigence minimale de fonds propres, fondées sur des quantifications forfaitaires captant une hiérarchie des risques qualitativement patente mais sans prétention quantitative.

En termes d'évaluation du bilan, insistons sur le fait que les estimations sont sciemment biaisées : en non vie par exemple, les estimations de versements futurs qui constituent les provisions ne sont pas actualisées, ce qui conduit mécaniquement à les surestimer ; si les PPNA d'une branche sont trop faibles, on les corrigera d'une PREC et si celles d'une autre branche sont trop élevées, on n'en tiendra pas compte ; en vie, les tables de mortalité sur lesquelles s'appuient les provisionnements des rentes sont calibrées pour *sous-estimer* la mortalité afin d'accroître les provisions, alors que les tables de mortalité sur lesquelles s'appuient les provisionnement des temporaires décès sont différentes, et sont à l'inverse calibrées pour *surestimer* la mortalité, là encore afin d'accroître les provisions. A l'actif, si les actions et les obligations sont en moins-value latente globale, on le comptabilisera mais inversement, si elles sont en plus-value latente globale, on ne prendra pas cela en compte.

En ce sens, la réglementation n'offre pas de cohérence de lecture, de méthode de calcul : les provisions ne sont pas actualisées en non vie alors qu'elles le sont en vie, deux engagements envers une même personne seront calculés à l'aide

d'hypothèses de mortalité différentes. En revanche, elle propose une cohérence d'objectif : systématiquement, la distorsion est telle que la vision est artificiellement négative ; on sait donc que la situation réelle étant plus favorable, l'assureur sera en mesure d'indemniser les assurés. Il apparaît ainsi que les règles d'évaluation du bilan d'une société d'assurance ont été, dans le cadre de Solvabilité 1, conçues dans un objectif de protection des assurés et non de communication financière.⁹⁴

1.4 Solvabilité 2⁹⁵

Solvabilité 2 fait voler en éclats ces principes pour leur substituer un système fondamentalement différent. Cette transition peut être symbolisée par un échange ayant eu lieu entre deux régulateurs lors de la conception de Solvabilité 2, relaté par un des protagonistes :

« Il [le représentant de l'autorité de contrôle néerlandaise] me disait que quand il contrôlait une entreprise, il voulait connaître sa vraie valeur. Il m'a pris pour un fou quand je [le représentant de l'autorité de contrôle française] lui ai dit que même en supposant qu'elle existe, moi je m'en foutais de la vraie valeur. Mais c'est vrai, je m'en foutais : la seule chose qui était importante pour moi, c'était que le biais soit dans le bon sens ! » (Entretien)

Dans ce cadre, comment Solvabilité 2 a-t-elle conçu et défini la « vraie » valeur d'un organisme d'assurance ; comment l'exploite-t-elle à des fins prudentielles de protection des assurés ?

⁹⁴ Une critique, non recevable à mes yeux, est parfois adressée : dans la marge disponible, les plus-values latentes étaient elles aussi prises en compte, au-delà de l'écart entre actifs et provisions. Certains présentent ceci comme un biais conduisant à surestimer les fonds propres disponibles, en s'appuyant sur le fait que les plus-values latentes, lorsqu'elles seront réalisées, devront pour partie être reversées aux assurés *via* la participation aux bénéfices, et qu'on ne peut donc pas les considérer intégralement comme représentant la richesse de l'entreprise. Nous ne partageons pas ce point de vue. De fait, dans une logique prudentielle, la valorisation n'est utile qu'en cas de difficultés de l'entreprise, lorsque sa survie est en jeu. Dans un tel cas de figure, l'entreprise ne réalisera pas de bénéfices et ne partagera donc rien avec les assurés. Les plus-values latentes, dans le cadre d'une comptabilité prudentielle, constituent donc bien une équivalence de fonds propres.

⁹⁵ Directive 2009/138.

Actifs

Le principe sous-jacent de Solvabilité 2 est celui de la justesse des évaluations des différents postes du bilan. Par « juste », on entend « non biaisé », non déformé par une convention, « objectif ». Pour les actifs, il s'agira donc de la *valeur de marché*, qui reflète le montant que l'assureur peut obtenir en échange du titre s'il le réalise au moment où son bilan est évalué.

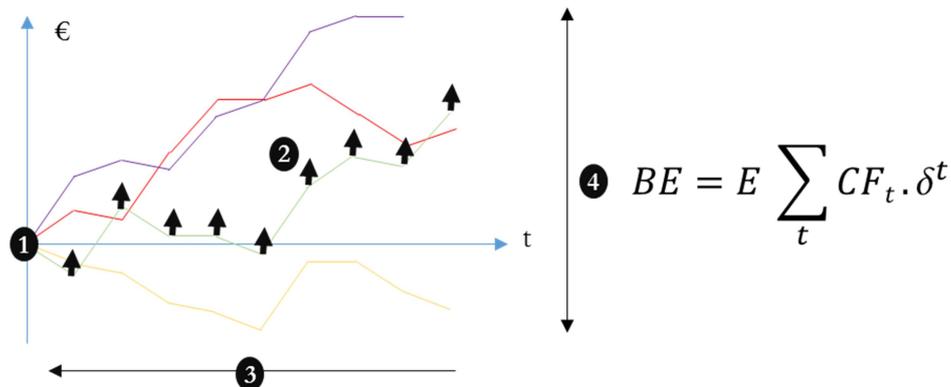
Provisions

De même que pour les actifs, les provisions vont être sous Solvabilité 2 évaluées en « valeur de marché ». Toutefois, étant donné qu'il n'existe pas de marché des passifs d'assurance, cette valeur ne peut être obtenue par simple constatation d'un prix de transfert. Elle sera donc déterminée à partir de modèles, reconstruite selon des règles de telle sorte qu'elle corresponde à ce que serait une valeur d'échange. Ce montant correspond à la somme de deux composantes : le *best estimate* (BE) et la marge pour risque (MVM, *market value margin*).

Le BE est la « meilleure estimation » de la valeur actuelle des versements futurs réalisés par l'assureur à l'assuré. La grandeur mathématique correspondant à cette notion de « meilleure estimation » n'est pas formellement définie. En pratique, le consensus tacite est de considérer qu'il s'agit de l'espérance de ces flux futurs actualisés. Pour le calculer, il faudra faire appel, en non-vie, à des techniques actuarielles classiques, comparables à celles utilisées en Solvabilité 1, mais sans marge de prudence ; en vie, étant donné que le taux de rendement servi aux assurés incorpore une partie des bénéfices réalisés par l'entreprise (contrats dit « *with profits* », cf. 1.1.2), bénéfices qui sont contingents aux évolutions futures des marchés financiers, il sera nécessaire de simuler ces évolutions pour déterminer leur espérance et partant le montant du BE⁹⁶. Il est nécessaire pour cela d'utiliser (i) des outils de projection stochastiques des

⁹⁶ En effet, en raison des fortes non-linéarités existant entre les évolutions des marchés et le taux servi aux assurés, l'espérance du taux servi sur l'ensemble des scénarii n'est pas le taux servi sur l'espérance de l'ensemble des scénarii.

évolutions des marchés financiers (ESG, *Economic Scenario Generator*), (ii) des outils de modélisation des interactions actif/passif, (iii) la fixation de nombreuses conventions de calcul partagées (e.g. le niveau des taux de référence), et (iv) le calibrage ou choix d'un très grand nombre d'hypothèses internes à l'entreprise (e.g. les *management rules* qui représentent la façon dont le management de l'entreprise interagira dans le futur avec les évolutions des marchés et les comportements des clients).

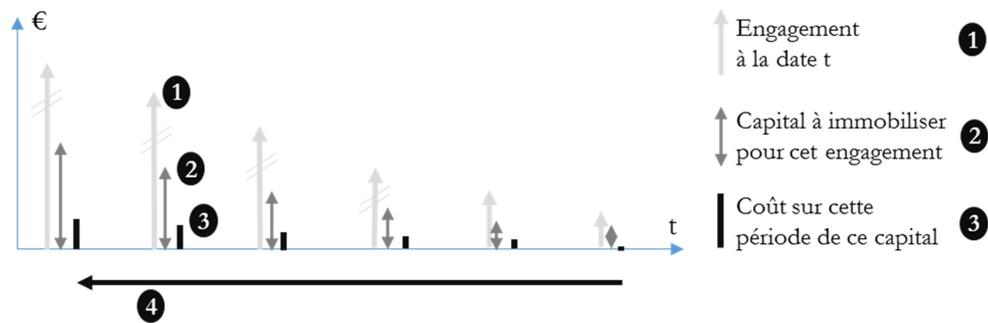


- ① Simuler de nombreux scénarii d'évolution des marchés financiers
- ② Pour chaque scénario, déterminer les cash-flow versés à l'assuré chaque année : CF_t
- ③ Pour chaque scénario, déterminer la valeur actualisée de l'ensemble de ces flux : $\sum_t CF_t \cdot \delta^t$
- ④ Calculer la moyenne sur l'ensemble des scénarios

Schéma 2 : détermination du best estimate

La *MVM* correspond au coût de portage des engagements. Le principe sous-jacent est le suivant : Supposons qu'un assureur A veuille transférer un portefeuille de contrats à un assureur B, quel montant d'actifs doit-il lui fournir pour que B accepte de reprendre les engagements associés au portefeuille ? B devra, en espérance, verser un certain montant aux assurés. La valeur actuelle de ce montant est le BE, que nous venons déjà de décrire. Mais B devra également, du fait de cette activité, immobiliser des capitaux : les fonds propres que le régulateur exige qu'il détienne parce qu'il porte des engagements dans le cadre de son activité d'assurance, le matelas de sécurité qui doit lui permettre d'absorber les chocs. Or, une immobilisation de capital représente un coût (il

donne lieu à des intérêts s'il s'agit d'emprunt subordonné, à des attentes de dividendes s'il s'agit de l'émission d'actions). Ce coût est la MVM.⁹⁷



- ① Déterminer, à chaque année dans le futur, quels seront les engagements résiduels
- ② En déduire, pour chaque année, quelle sera l'exigence de capital associée à cet engagement : K_t
- ③ En obtenir, pour chaque année, le coût correspondant : $CoC \cdot K_t$ (e.g. $4\% \cdot K_t$)
- ④ Déterminer la valeur actuelle de ces coûts cumulés : $MVM = \sum_t CoC \cdot K_t \cdot \delta^t$

Schéma 3 : principe de détermination de la MVM

Ainsi, lorsque A cède son portefeuille à B, le montant des actifs à transférer en face correspond à la somme du BE (qui sera reversé aux assurés) et du coût du capital à immobiliser, la MVM : nous avons pu construire la « valeur de marché » d'une provision.

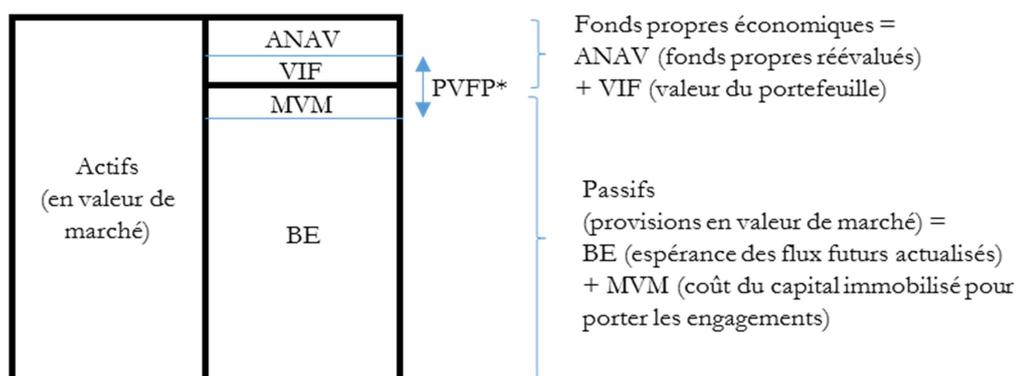
Fonds propres disponibles

Les fonds propres disponibles seront également déterminés en vision dite « économique » : ils contiennent l'appréciation des actifs dans lesquels la mise de départ des actionnaires a été investie, ainsi que la valeur des contrats en portefeuille.

⁹⁷ Pour l'estimer, on projette l'écoulement des provisions dans le temps, on en déduit le déroulé du montant de l'exigence de fonds propres dans le temps, on applique un pourcentage correspondant au taux de rémunération des fonds propres (le *CoC*, *cost of capital*, aujourd'hui fixé à 6%), puis on actualise le déroulé des coûts annuels qui en résultent.

De fait, correspondant à l'écart entre l'actif et le passif, ils doivent être évalués en cohérence avec les nouveaux principes retenus. Techniquement, ils correspondent à la somme de deux composantes : l'*Adjusted Net Asset Value* (ANAV), et la *Value of In-Force* (VIF). L'ANAV correspond à la définition historique des fonds propres, celle de Solvabilité 1, mais revalorisée des plus ou moins-values latentes des actifs situés en face puisque nous raisonnons désormais en valeur de marché. La VIF représente la valeur du portefeuille : elle correspond aux perspectives de profits futurs⁹⁸, mais corrigées des coûts de portage du capital, la MVM présentée précédemment.

Ceci permet, en théorie⁹⁹, d'obtenir le bouclage requis entre chacun des postes du bilan tel que présenté ci-dessous.



* Profits futurs (PVFP), se décomposant entre la rémunération des capitaux immobilisés (MVM) et la valeur résiduelle du portefeuille (VIF)

Schéma 4 : vision schématique d'un bilan Solvabilité 2

Exigences de capital

Une entreprise est en faillite si le montant de ses actifs devient inférieur au montant de ses passifs. Solvabilité 2 fait reposer l'exigence de capital, *Solvency*

⁹⁸ *Present value of future profits (PFVP)*, qui est à l'actionnaire ce que le BE est à l'assuré : une actualisation des flux futurs.

⁹⁹ En pratique, la complexité des calculs de PVFP et de BE est telle qu'il y a des « fuites », parfois non négligeables, nécessitant le recours à des clés de répartition de l'excès ou du déficit global sur les différents postes

capital Requirement (SCR)¹⁰⁰ sur une mesure de risque, la *value-at-risk* (VaR) des fonds propres ainsi définis à 1 an à 99,5%. En d'autres termes, le matelas doit être suffisamment épais pour que la probabilité qu'il devienne négatif, donc que l'entreprise soit en faillite à horizon un an, ne soit que de 0,5%.

D'un point de vue méthodologique, la réglementation élabore une cartographie des risques (e.g. baisse des marchés action, survenance d'une catastrophe naturelle, dérive de la longévité, etc.) et demande d'évaluer, en cas de survenance d'un événement rare d'une amplitude correspondant à cette probabilité, l'impact sur les fonds propres économiques. L'ensemble de ces impacts sont ensuite agrégés et, après prise en compte de plusieurs correctifs (non survenance simultanée de tous les événements adverses, non linéarités des mécanismes tels que la participation aux bénéfices ou l'absorption partielle des pertes par la réduction des impôts sur profits futurs en cas d'absence d'obtention de ces profits), on obtient l'impact sur les fonds propres correspondant à la survenance d'un événement bicentenal global : il s'agit de l'exigence de capital.

¹⁰⁰ La réglementation introduit deux seuils : le SCR évoqué *supra*, et le *minimum capital requirement* (MCR), environ trois fois plus faible. Le MCR est déterminé de façon plus rustique, à l'instar des exigences de fonds propres de Solvabilité 1, mais ne peut être à l'extérieur d'un tunnel dont le plancher et le plafond sont un pourcentage du SCR. Formellement, couvrir le MCR est une condition d'exercice de l'activité, alors que le non-respect du SCR requiert uniquement la mise en œuvre d'un plan d'amélioration du ratio de solvabilité. Dans les faits, notamment pour des raisons de communication financière, réglementaire et commerciale, seul le SCR fait l'objet d'attention et est considéré par les acteurs de l'industrie comme la véritable exigence de fonds propres.

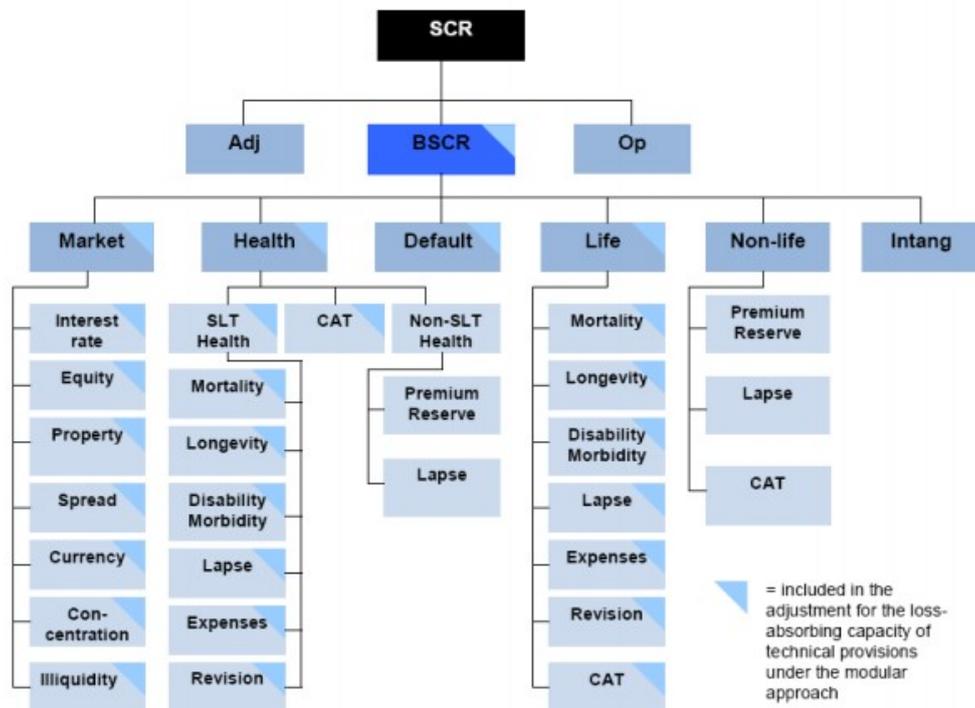


Schéma 5 : Cartographie des risques modélisés sous Solvabilité 2

(source : ACPR)

Ce montant peut être déterminé soit par l'utilisation d'une formule standard, dont l'architecture générale d'agrégation décrite *supra* et l'amplitude des stress sont fixés par la réglementation¹⁰¹ ; soit par le recours à l'architecture de la formule standard, mais avec des calibrages de chocs calibrés sur le portefeuille de l'entreprise (USP, *undertaking specific parameters*) ; soit par un modèle interne laissant entière liberté à l'entreprise, en termes de méthodologie mathématique, pour déterminer le quantile global des fonds propres à 99,5% à 1 an. Dans les deux derniers cas de figure, l'approbation de l'autorité de contrôle est requise. L'ampleur du recours à ces options est très variable selon les marchés nationaux.

¹⁰¹ En vie, le calcul du montant nécessite quoi qu'il en soit le développement d'un modèle d'interaction actif passif et le recours à des projections d'actifs, puisque ceci est nécessaire à la détermination du bilan prudentiel qui est choqué.

Synthèse

L'architecture des exigences quantitatives de Solvabilité 2 articule fonds propres économiques, provisions en valeur de marché et exigences de capital, tous fondés sur la simulation de flux stochastiques futurs : chaque grandeur est déterminée à partir de modèles mathématiques complexes, qui réconcilient *in fine* avec la valeur de marché des actifs. Elle est à ce titre remarquable de cohérence interne.

La valorisation des provisions bascule d'une logique d'indemnisation (des provisions « suffisantes », « suffisamment prudentes ») à une logique de transfert (en « juste valeur »). D'un point de vue opérationnel, ceci n'est *a priori* pas contradictoire avec une finalité prudentielle. De fait, l'idée sous-jacente est que, si les fonds propres de l'assureur devenaient insuffisants pour absorber d'éventuels chocs futurs, alors l'autorité de contrôle pourrait procéder au retrait d'agrément et au transfert de son portefeuille vers un autre assureur : avec des provisions évaluées en valeur de marché et des actifs en face en valeur de marché, il serait possible de trouver un repreneur et les intérêts des assurés seraient entièrement préservés. Toutefois, d'un point de vue conceptuel, on voit ici s'opérer un glissement important, puisque cette évolution traduit le passage d'un critère relatif aux assurés (des provisions suffisamment prudentes pour couvrir les engagements) à un critère relatif aux marchés financiers (des provisions correspondant à la valeur d'échange d'un passif d'assurance si un marché existait pour ces titres).

Qui plus est, Solvabilité 2 offre une grille d'analyse particulièrement puissante en ce sens qu'elle s'affranchit des contingences liées au contrat : là où nous avons dû décrire, pour Solvabilité 1, le provisionnement non vie d'une part et le provisionnement vie d'autre part ; là où, en non vie par exemple, telle spécificité contractuelle nécessitait tel type de provision supplémentaire (la *provision pour risque croissant*, déjà évoquée, pour des contrats de prévoyance sur plusieurs années, ou encore la *provision pour sinistres non encore manifestés* en assurance construction), ici, avec le concept de *best estimate* et de *marge pour*

risque, nous avons une vision homogène valable quelles que soient les branches d'assurance considérées, quelles que soient les spécificités produit de chaque marché national. Solvabilité 1 était ancré dans les spécificités techniques de l'industrie considérée, les épousant une à une ; avec Solvabilité 2 en revanche, le bilan repose sur des concepts suffisamment théoriques pour être unifiés et peut donc être lu par des analystes n'ayant aucune expérience particulière de l'activité considérée.

Enfin, Solvabilité 2 marque le passage d'une exigence de capital largement forfaitaire, (i) traduisant de façon qualitative une hiérarchie des risques des activités d'assurance et (ii) indépendante des prises de risque à l'actif qui étaient encadrées par des contraintes de dispersion¹⁰², à une exigence de capital ayant pour ambition de refléter finement, de façon quantifiée, la prise de risque de l'entité.

Nous voyons apparaître, à travers ces trois changements majeurs, à quel point Solvabilité 2 rend totalement imbriqué son objectif initial de protection des assurés (qu'elle est réputée partager avec Solvabilité 1) avec l'ambition de développement d'un marché intérieur européen *des passifs d'assurance*¹⁰³ (objectif qui était totalement orthogonal à l'ambition de Solvabilité 1). De fait, Solvabilité 2 mentionne explicitement l'importance d'une « *allocation efficace des capitaux dans l'Union Européenne* » pour « *la protection des preneurs [d'assurance]* » (considérant 14). Ces deux objectifs peuvent-ils être poursuivis simultanément ? La grille d'analyse pertinente du point de vue des choix d'investissement des analystes financiers est-elle la même que celle qui permet de protéger les preneurs d'assurance ? Si tel n'était pas le cas, les choix retenus ont-ils conduit à subordonner la création d'un marché des portefeuilles

¹⁰² Règles en pratiques suffisamment souples pour n'être jamais saturées, sauf cas spécifiques où la stratégie d'investissement interfère avec la structure capitalistique d'un groupe ou bien entités de taille particulièrement petite.

¹⁰³ Formellement, la promotion d'un marché intérieur apparaît très explicitement dans les considérants de Solvabilité 2 (considérants 3 et 11), mais cela vise *a priori les produits d'assurance* et non les *portefeuilles d'assurés*.

d'assurance à la « protection des preneurs », ou bien à renverser les priorités et faire primer le moyen sur la fin ?

2. Pour qui les concepts de Solvabilité 2 sont-ils signifiants ?

Afin de répondre à cette question, nous analyserons successivement les choix conceptuels liés à chacun des principaux postes du bilan :

- le *best estimate* (en nous concentrant ici sur les assureurs vie),
- puis le montant de provision dans sa globalité (en prenant en compte la MVM),
- puis le bouclage entre la marge disponible qui en résulte et l'exigence de capital qui en est déduite,

avant de prendre du recul pour apprécier ces éléments au sein du système quantitatif appréhendé dans son ensemble.

Nous analyserons cette question en mettant en regard le point de vue de deux catégories d'acteurs :

- d'une part, la collectivité des assurés, représentés comme nous l'avions indiqué par le régulateur ;
- et d'autre part, les analystes financiers. Par cette catégorie, nous ne distinguons pas à ce stade les analystes extérieurs, représentant les actionnaires cherchant à valoriser ou s'informer sur la valeur d'une entreprise d'assurance dans un objectif de comparaison des unes par rapport aux autres, des analystes internes aux holdings d'assurance, cherchant à valoriser une ligne d'activité, un portefeuille d'assurés, un segment de passifs, afin d'apprécier (en relatif de façon générale, dans l'absolu lorsque nous le préciserons) s'il serait opportun de le céder ou de l'acheter. Il est légitime à ce stade de ne pas distinguer ces deux

objectifs car la vision retenue par Solvabilité 2 assimile la valorisation d'une entreprise à la somme des valorisations de ses composantes¹⁰⁴.

2.1 En vie, un *best estimate* ayant du sens pour l'actionnaire, non pour l'assuré

Les provisions d'assurance vie représentent l'essentiel du bilan des assureurs. Rappelons que, dans ces branches, les assureurs, lorsque les revenus de leurs placements sont supérieurs au taux de rendement qu'ils ont garanti à leurs assurés, doivent reverser à ceux-ci une proportion minimale (typiquement 90%) des bénéfices. Le montant des bénéfices qui seront ainsi versés aux assurés au-delà des minimums garantis dépend donc de l'évolution à venir des marchés financiers, ainsi que des choix financiers (stratégie d'investissement) et commerciaux (verser une proportion plus importante que le minimum) de l'entreprise, et des interactions entre chacun de ces paramètres. Il s'agit d'éléments contingents à des phénomènes futurs, inconnus.

Le provisionnement sous Solvabilité 1 prend en compte de façon statistiquement prudente (c'est-à-dire avec des taux d'actualisation, de rendement ou de mortalité défavorables) uniquement les montants que l'assureur est *certain*¹⁰⁵ de devoir aux assurés :

Les provisions techniques d'assurance vie doivent être calculées selon une méthode actuarielle prospective suffisamment prudente, tenant compte de toutes les obligations futures conformément aux conditions établies pour chaque contrat en cours, et notamment:

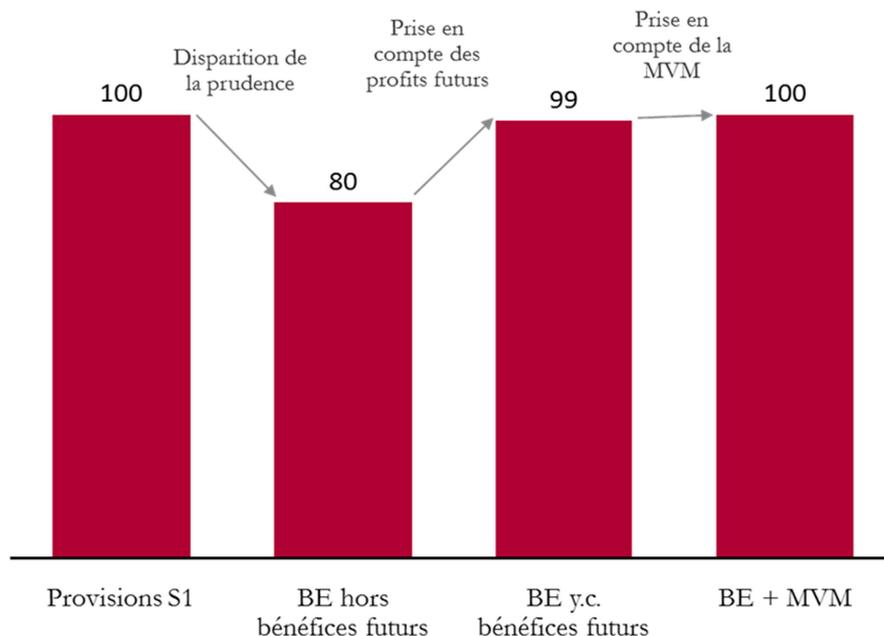
- de toutes les prestations garanties, [...]
- des participations aux bénéfices auxquels les assurés ont *déjà* (je souligne) collectivement ou individuellement droit, [...], [donc déjà garantis]

Sous Solvabilité 2 en revanche, le *best estimate* prend en compte de façon non prudente (c'est-à-dire en préférant ne pas biaiser l'appréciation quitte à avoir une fois sur deux une mauvaise surprise), les montants que l'assureur versera

¹⁰⁴ Nous négligeons les effets de troisième ordre ré-incorporant *via* la MVM certains des bénéfices de diversification dans les provisions.

¹⁰⁵ Il s'agit d'une certitude statistique, puisqu'on prend en compte la probabilité de décès en retraite par exemple

aux assurés en incluant l'espérance actualisée des montants liés à ces perspectives de profits futurs.



Provisions S1 : provisions prudentes (e.g. mortalité sous-estimée pour les engagements de rente)

BE hors bénéfices futurs : provisions sans marge de prudence, et où seuls sont pris en compte les engagements minimums garantis de l'assureur

BE y.c. bénéfices futurs : provisions sans marge de prudence, où les perspectives de partage des bénéfices reversés aux assurés sont pris en compte

BE+MVM : provisions S2, sans marge de prudence, avec perspectives de profits futurs reversés aux assurés, et avec la MVM (coût actualisé d'immobilisation du capital)

Schéma 6 : visualisation des montants relatifs de provision en vie¹⁰⁶¹⁰⁷

¹⁰⁶ Source : pour les trois premières colonnes : rapport QIS 1 (EIOPA, 2006) moyenne pondérée sur les marchés dont les données ont été fournies par l'EIOPA et partagées par celle-ci. Pour la dernière colonne : reconstruite à partir de la précédente en appliquant la MVM médiane fournie dans le rapport QIS 4 (EIOPA, 2008). Nous avons retenu la médiane car la moyenne semble faussée par des valeurs extrêmes.

¹⁰⁷ **Le fait que le montant total sous Solvabilité 2 corresponde presque exactement au montant total sous Solvabilité 1 ne doit pas amener à tirer de conclusion sur cette équivalence**, d'une part en raison de la très forte volatilité de ces chiffres, telle qu'une mesure à une autre date aurait conduit à des résultats sensiblement différents et, d'autre part, car l'évolution de l'évaluation des provisions ne peut être appréciée qu'en regard de l'évaluation des autres postes du bilan et de l'exigence de marge.

On constate sur le graphique ci-dessus que la prise en compte des profits futurs a un impact considérable sur le montant des provisions. En effet, ceci représente, un montant de l'ordre de 20 à 25 % (dans le cas d'espèce passage d'un montant de 80 à 99) là où, à titre de comparaison, l'exigence de fonds propres d'une compagnie d'assurance vie sous Solvabilité 1 était de l'ordre de 4 à 5% des provisions. Dans quelle mesure ce changement de définition accroît-il ou dégrade-t-il la pertinence de l'information fournie ?

L'objectif informationnel n'est plus la protection de l'assuré contre la faillite

La finalité d'une régulation prudentielle est de s'assurer que les engagements pris par les organismes envers les assurés seront tenus. Dans ce cadre, est-il utile de prendre en compte parmi les engagements l'espérance des profits futurs qui, s'ils sont réalisés, devront être versés aux assurés ?

De deux choses l'une : soit les investissements de l'entreprise procureront des rendements trop faibles pour servir les taux minimums garantis aux assurés, soit les taux de rendement des actifs seront suffisants pour lui permettre de servir ces taux, et au-delà de réaliser des bénéfices. Dans le premier cas de figure, l'entreprise ne réalisera pas de bénéfices : la protection des assurés ne passe donc pas par le fait d'évaluer combien il leur serait versé en moyenne si l'entreprise en réalisait. Seul l'engagement minimal (le capital garanti, les participations *déjà* acquises) devra être versé : c'est à l'évaluation de *cet* engagement, des actifs en représentation et du matelas de sécurité constitué par les fonds propres qu'il convient de s'attacher pour s'assurer que les engagements pris à l'égard des assurés seront tenus. Dans le second cas de figure, l'assureur réalisera des bénéfices, au-delà des coûts que représente l'exécution des engagements minimaux. Dès lors, il n'y a pas à craindre de faillite de sa part : les engagements envers les assurés seront tenus, et ceux-ci percevront, au fur et à mesure du dégagement des bénéfices, leur participation. Autrement dit, soit l'organisme est en bonne santé (rendements suffisants ou élevés), et une réglementation prudentielle est inutile, soit l'organisme est en mauvaise santé (rendements insuffisants ou pertes), et la non prise en compte

de la participation aux bénéfices ne conduit pas à sous-estimer les engagements envers les assurés.

Ainsi, nous voyons que valoriser, dans les engagements de l'assureur envers les assurés, l'espérance des profits que le premier versera aux seconds au cas où les marchés se portent bien, n'apporte rien en termes de protection des assurés.

En revanche, il s'agit là d'une information utile à l'actionnaire. De fait, si celui-ci souhaite savoir quelles sont ses perspectives de profits futurs, il ne peut se contenter d'une estimation des profits bruts réalisés par l'entreprise grâce à ses placements : il a également besoin de savoir quelle part de ces profits sera versée aux assurés, afin d'évaluer les perspectives de profits nets.

L'estimation est porteuse de sens pour l'actionnaire et non pour l'assuré

Toutefois, une autre justification peut être envisagée. L'actif étant évalué en valeur de marché, il est réputé valoriser les perspectives de rendements futurs. Afin d'avoir un bilan cohérent, il est donc nécessaire de pouvoir répartir ces résultats futurs entre les fonds propres (pour l'assureur) et les passifs (pour les assurés). On pourrait donc se dire que, à défaut d'être utile en termes de protection des assurés, une estimation des revenus futurs et de leur répartition serait source d'information pour les assurés. Est-ce le cas ?

Frezal (2015, a), établit le champ de pertinence des indicateurs statistiques. Rappelons ses conclusions. Lorsqu'un événement aléatoire se réalisera plusieurs fois (situation dite *hétérogénéité*), les statistiques sont porteuses de sens ; par exemple, lorsque l'Etat ou un casino organise des jeux de hasard, même si chaque gain individuel est aléatoire, ils s'incarneront collectivement dans une moyenne : c'est en s'appuyant sur cette grandeur que l'institution pourra développer des jeux structurellement rentables. L'espérance a donc une signification tangible. En revanche, lorsqu'un événement aléatoire ne se réalisera qu'une unique fois (situation dite *aléa*), les grandeurs statistiques ne sont pas porteuses de sens : si on me propose de jouer une fois et une seule à un jeu où je gagne ou perd 1M€, ma réponse dépendra sans doute peu du fait que la probabilité de gagner soit d'une chance sur quatre ou de trois chances

sur quatre, donc dépendra peu de mon espérance de gain (ou de façon duale, si on me propose de jouer une fois et une seule à pile ou face, avec une mise de 1 M€, ma décision sera sans doute indépendante du fait qu'en cas de victoire, le gain soit de 1 ou 3 M€, donc indépendante de l'espérance).¹⁰⁸

Dans le cas présent, l'assuré est généralement en situation aléa : il ne diversifie pas ses assurances vie entre de nombreux assureurs. Ses profits futurs dépendront de l'intersection entre la stratégie d'investissement de l'assureur et l'évolution des marchés financiers, intersection dont il n'observera le résultat qu'une fois sur son horizon d'investissement. Ainsi, son espérance de profits futurs ne s'incarnera jamais dans une moyenne (Frezal, 2015, b) : l'espérance de ses profits futurs n'est pas une estimation de ses profits futurs effectifs et ne lui apporte pas d'information sur laquelle il lui serait pertinent de fonder une décision.

A contrario, l'actionnaire peut investir dans des titres d'assureurs ayant des stratégies d'investissement différentes, et surtout investir sur des titres et catégories d'actifs en dehors du secteur de l'assurance (actions de compagnies ou titrisation de portefeuilles d'assurances) : l'actionnaire est en situation hétérogénéité. Par conséquent, ses espérances de profits futurs s'incarneront globalement, par la loi des grands nombres, dans leur moyenne : cette espérance est porteuse de sens pour lui (a minima d'un point de vue relatif, en termes de comparaison entre deux assureurs ou deux portefeuilles).

Nous voyons donc que le principe même du *best estimate* (et de son complémentaire que sont les fonds propres économiques) est un principe qui, en tant qu'espérance, a un sens opérationnel pour les actionnaires (qui investissent dans un portefeuille diversifié dont les acteurs de l'industrie de l'assurance ne représentent qu'une part de l'allocation), mais qu'il n'a pas de

¹⁰⁸ Cette position est cohérente avec les heuristiques comptables usuelles : de tout temps en non vie, les assureurs ont provisionné statistiquement les provisions. De fait, ils sont confrontés à de nombreux sinistres et la moyenne est tangible. En revanche, lorsqu'on est confronté à un nombre d'engagements limité (par exemple quand on comptabilise ses propres dettes), c'est-à-dire en situation aléa, on ne comptabilise généralement pas l'espérance de l'engagement, mais sa valeur nominale.

signification (autre que totalement virtuelle), pour un assuré qui sera confronté à une seule réalisation de rendement et de taux servi en résultant.

L'arbitrage biais/marge d'erreur correspond à une logique d'investissement et non de régulation

Le passage d'un bilan fondé sur des provisions sociales à un bilan fondé sur le *best estimate*, cette bascule d'une évaluation prudente des engagements fermes vers une évaluation « juste » des dettes probables, c'est-à-dire contingentes à la réalisation de scénarii économiques non définis et non identifiés, peut être lue comme une évolution de l'arbitrage fait par les normalisateurs entre biais et marge d'erreur.

Les comptes d'une société d'assurance sont des indicateurs déterminés, du fait de l'activité de l'organisme, par des estimateurs statistiques. Un indicateur et un estimateur sont notamment caractérisés par leur biais et leur marge d'erreur, deux grandeurs dont l'optimisation peut nécessiter un arbitrage¹⁰⁹. Dans le cas de Solvabilité 1, l'indicateur des engagements est doublement biaisé : d'une part, il sous-estime le montant qui serait statistiquement versé en ne prenant pas en compte l'espérance de participation aux bénéfices : nous avons vu que d'un point de vue prudentiel, pour la protection des assurés, cela ne portait pas à conséquence de se concentrer uniquement sur les engagements fermes, non contingents aux résultats financiers ; d'autre part, il surestime le montant d'engagements fermes (principe de prudence dans le choix des taux de rendement, des tables de mortalité, etc.), ce qui correspond à un objectif prudentiel visant à garantir que les engagements fermes envers les assurés

¹⁰⁹ Par exemple, en termes d'estimateurs statistiques, la correction du biais de non réponse par une repondération des réponses peut générer un accroissement de la variance. En termes d'indicateur, un bon exemple est celui de la recommandation formulée par Gary Becker dans son discours de réception du prix Nobel : intégrer le capital humain dans le bilan des entreprises. En effet, ne pas comptabiliser le capital humain conduit à sous-estimer les actifs de l'entreprise, et introduit donc un biais que sa préconisation supprime ; mais cela accroît considérablement la marge d'erreur : quelle était la valorisation d'un Steve Jobs pour Apple par exemple, 10 M\$ ou 100Mds\$? La réduction du biais ne se ferait qu'au prix d'un accroissement considérable de la marge d'erreur, susceptible de noyer l'information dans le bruit. Sa préconisation n'a de fait pas été retenue.

seront tenus. Le corollaire de ces choix est que l'estimateur (biais compris) est précis : la plupart des éléments inconnus (taux de mortalité précis, rendement des marchés financiers, etc.) ont été éradiqués des calculs pour être remplacés chacun par une hypothèse précise commune et partant connue de tous ceux qui sont familiers de la comptabilité considérée.

Solvabilité 2 supprime ces biais, puisqu'elle demande d'une part, de prendre en compte les perspectives de participation aux bénéfices au-delà de l'engagement ferme et, d'autre part, de ne pas faire preuve de prudence. Quelles en sont les conséquences ? Le premier impact est, en supprimant la prudence, de supprimer un biais dans l'évaluation. Qu'en est-il de l'impact sur la marge d'erreur ? D'une part, l'estimation de ces perspectives de bénéfices est imprécise, puisque ces bénéfices sont, contrairement aux engagements fermes, contingents à l'évolution future inconnue des marchés financiers. D'autre part, la substitution d'une justesse à la prudence nécessite de modéliser au plus près de chaque entreprise et de chaque scénario financier envisageable. Chaque organisme va élaborer *son* jeu de scénarii économiques (au sein d'un Groupe, ils pourront différer d'une filiale à l'autre) variant suivant la mode *du moment*, les comportements de *ses* assurés, les choix de *ses* dirigeants. La méthode de calcul devient donc imprécise. On introduit de ce fait une dispersion dans l'estimation, une double marge d'erreur : d'une part, parce qu'on fait des projections futures, sur ce qui sera versé aux assurés et, d'autre part, parce que l'ambition de justesse requiert une « ad-hocité » de la méthode de calcul. Ainsi le passage global d'une provision en Solvabilité 1 à une provision en Solvabilité 2 marque l'évolution d'un arbitrage entre biais et marge d'erreur : un biais plus faible, et une marge d'erreur plus élevée. Cette évolution va l'encontre des pratiques actuelles des statisticiens qui, après s'être efforcés de construire des estimateurs minimisant le biais, cherchent actuellement des estimateurs où le

biais peut être accru si cela permet de réduire la variance¹¹⁰. Comment interpréter une telle évolution ?

Faisons une analogie : ce cas est analogue à celui où on déciderait de demander pour les papiers d'identité la hauteur d'une personne non plus entre la plante des pieds et le haut de sa tête, mais entre le bas des talons de chaussure et le haut de sa tête. Chaque répondant interprètera différemment la prise en compte du talon de la chaussure : deux femmes qui sont toute la semaine en escarpins et tous les week-ends en ballerines vont choisir, pour l'une qui voyage pour affaires, la hauteur des escarpins et pour l'autre qui voyage souvent le week-end, la hauteur des ballerines. Ainsi, cette nouvelle métrique réduit le biais (la hauteur du talon de chaussure était auparavant toujours négligée, comme si elle était nulle), mais augmente la marge d'erreur (la précision de mesure habituelle de la taille d'un individu est de l'ordre du centimètre, alors que l'imprécision liée au fait qu'on peut ne pas porter ses chaussures habituelles sera de l'ordre de la dizaine de centimètres). Cette nouvelle métrique est-elle plus pertinente ou moins pertinente que l'ancienne ? Cela dépend pour qui. Pour la personne de la police de l'air et des frontières qui, assis dans sa guérite, voit défiler les passagers debout derrière le comptoir, droits dans leurs chaussures, la seconde métrique permet de mieux juger, d'un coup d'œil, si la personne face à lui a une taille correspondant à la pièce d'identité : cette seconde métrique est plus efficace, donc plus pertinente. À l'inverse, face à un suspect donné, dans le cas d'un contrôle plus poussé, pour le policier pouvant passer davantage de temps à faire des mesures, la première métrique est sans doute plus fiable, puisqu'indépendante du fait qu'on ne porte pas ses chaussures habituelles, donc plus pertinente. Dans notre cas d'espèce,

¹¹⁰ En notant B le biais et V la variance, les estimateurs aujourd'hui considérés comme préférables sont ceux qui minimisent B^2+V et non ceux qui minimisent B .

On pourra également noter que, d'un point de vue statistique, un biais est toujours défini relativement à un modèle donné, et que qualifier un biais de biais suppose que le modèle sous-jacent dont on cherche à estimer les paramètres soit juste.

celui de la régulation, aux yeux de quels acteurs cette évolution entre biais et marge d'erreur améliore-t-elle et dégrade-t-elle la pertinence de la métrique ?

Nous avons vu que, en offrant un cadre transversal théorisé, affranchi du lien avec les spécificités des produits et des marchés d'assurance donnés, Solvabilité 2 permettait de gagner en comparabilité entre pays, grâce à un unique concept s'appliquant à toutes les activités. Elle peut être ainsi lue comme une mise sur pied d'égalité entre les différents analystes, puisqu'elle fait disparaître l'avantage comparatif lié à la connaissance fine d'un marché national : avec notre réforme, puisque la hauteur du talon est déjà statistiquement prise en compte dans la taille figurant sur la pièce d'identité, notre représentant de la police de l'air et des frontières n'a plus besoin d'être expérimenté et de savoir, pour corriger son estimation de la taille, quel type de maintien des épaules est associé à des escarpins plutôt qu'à des ballerines.

Il serait toutefois abusif d'en conclure qu'un bilan de type Solvabilité 2 est plus aisément interprétable pour un œil extérieur. En effet, en vie, le caractère mécanique du provisionnement sous Solvabilité 1 était tel que les biais pouvaient être connus et partagés par tous (on sait qu'il faut ajouter la hauteur du talon pour apprécier l'altitude du regard qui nous fait face ; on sait que notre appréciation est faussée par la hauteur de la coiffure). Sous Solvabilité 2 en revanche, du fait de la complexité accrue des méthodes de détermination des postes du bilan (projections stochastiques, hypothèses nombreuses à la main des compagnies, etc.), la maîtrise de la façon dont chaque chiffre a été élaboré devient quasi-impossible et on perd en comparabilité entre deux entreprises menant une même activité. A titre d'exemple, le CFRO du sous-groupe assurance d'un grand groupe bancaire français expliquait que leurs outils de calcul du bilan et des exigences de capital (pourtant en formule standard) nécessitaient :

« 16 fois l'Iliade et l'Odyssée en nombre de lignes de code, 12 000 pages de documentation lorsque l'*Encyclopedia Universalis* en fait 20 000, etc. »

L'accroissement de comparabilité est donc essentiellement formel : *si la conceptualisation accrue offre une transversalité au niveau européen dans l'appréciation de la santé financière des entreprises, le prix à payer est une dégradation locale de la compréhension des comptes*. Pour reprendre notre analogie, c'est comme si on devait désormais communiquer à notre représentant de la PAF un manuel où il trouverait exposé le fait que, en fonction du jour de la semaine, de l'horaire du vol, de la classe du billet d'avion et de la coiffure, il doit désormais corriger statistiquement de telle ou telle façon l'espérance de hauteur du talon intégrée dans la hauteur fournie sur la pièce d'identité. Ce nivellement par le bas peut être vu *comme la suppression d'une forme de « délit d'initié » national dans la capacité de lecture des comptes* (tout comme la prise en compte de la hauteur statistique du talon serait une suppression du « délit d'initié » favorisant le jugement de ceux qui savent apprécier le maintien pour corriger la non prise en compte du talon). S'il semble difficile de supposer une intentionnalité de ce nivellement par le bas (cf. *infra*), le fait qu'il soit consciemment accepté est en revanche une hypothèse solide. A titre d'exemple, on peut citer cette affirmation d'un CRO d'une des quatre grandes banques françaises, lors d'un entretien portant sur les modèles d'optimisation des allocations d'actifs et de calcul des exigences de capital (qui partagent la double caractéristique d'une interaction actif/passif et de simulations stochastiques de scénarii financiers) :

« Les modèles ne valent rien. Mais le plus important est de ne pas fausser la concurrence. Donc, si tout le monde les utilise, il n'y a pas de raison d'en changer. » [propos recueillis en 2014]

Nous voyons donc ici que cette substitution d'une marge d'erreur à un biais dans les estimateurs des différents postes du bilan est pertinente du point de vue des marchés financiers, mais non du point de vue des régulateurs d'assurance nationaux : du point de vue des marchés financiers, en réduisant le coût d'entrée dans la compréhension des comptes et en supprimant un biais qui n'est bien connu que des spécialistes de chaque marché, cela permet d'améliorer le *level playing field* entre les différents investisseurs, favorisant une

fluidité des marchés de portefeuilles ; en revanche, du point de vue des régulateurs qui maîtrisaient chacun les spécificités de leur marché, cela les contraint à s'appropriier les spécificités d'implémentation de chaque organisme (spécificités qui plus est mouvantes) ce qui supposerait un changement d'ordre de grandeur dans les ressources mobilisées.

Cette mise en perspective permet de mieux interpréter le *satisfecit* accordé en 2011 à Solvabilité 2 par la présidente de l'ICA¹¹¹ qui, reflétant une opinion largement partagée, indiquait : « *disposer de règles et de principes communs aide à juger de la solvabilité d'une compagnie, et ce quel que soit son pays. Et c'est une excellente chose* ». Qui peut et doit être à même de juger de la solvabilité d'une compagnie ? Ce n'est pas l'assuré, qui n'a pas de compétences techniques dans ce domaine et délègue cette mission au régulateur. Il s'agit donc soit du régulateur, soit des analystes financiers. Or les compagnies d'assurances sont soumises au contrôle de leur régulateur national, l'EIOPA (le régulateur européen) ayant essentiellement un rôle de coordination. Chaque régulateur, connaissant son marché, est donc *a priori* peu concerné par l'ambition de « juger de la solvabilité d'une compagnie quel que soit son pays »¹¹². Les acteurs pour qui cette ambition représente « une excellente chose » sont donc les analystes financiers, qu'ils soient au demeurant externes aux entreprises, ou bien au sein des *holdings* des groupes internationaux.

Synthèse : calculer le best estimate, une évolution « inutile et incertaine » d'un point de vue prudentiel

Le remplacement d'un bilan fondé sur des provisions sociales par un bilan fondé sur des provisions reposant sur un *best estimate* ne correspond donc pas tant à un objectif de protection des assurés, visant à garantir la tenue des

¹¹¹ Dionne, M, présidente de l'ICA, (*Institute of Canadian Actuaries*), dans *l'actuariat* n°2 (2011), p41.

¹¹² En matière de contrôle des groupes, l'enjeu partagé par différents régulateurs est d'éviter le double emploi de fonds propres. La mise sous contrôle de cet aspect ne nécessite cependant pas d'homogénéité de la définition des fonds propres.

engagements pris à leur égard, qu'à un objectif d'évaluation des profits futurs des actionnaires.

Premièrement, avec Solvabilité 2, l'ambition de la comptabilité prudentielle n'est plus de savoir combien d'actifs doivent être détenus pour rembourser les assurés, mais comment les actifs détenus seront distribués entre actionnaires et assurés.

Deuxièmement, cette répartition des profits futurs entre actionnaires et assurés est fondée sur une espérance qui est porteuse de sens pour l'actionnaire (il s'agit d'un estimateur sans biais de la valeur de l'entreprise, estimateur qui s'incarnera dans une moyenne pour un actionnaire diversifiant ses placements), mais n'est pas porteuse de sens pour l'assuré (étant placé en situation aléa, l'espérance est pour lui une grandeur virtuelle).

Troisièmement, la suppression des biais dans l'estimation des postes du bilan se fait au prix d'un accroissement de la marge d'erreur. Ce déplacement de l'arbitrage entre biais et précision (auquel il doit nécessairement être procédé lors du choix d'un estimateur statistique) peut être considéré comme un gain de pertinence pour la communauté des investisseurs mais est une perte de pertinence pour chaque autorité de contrôle.

Ainsi, il semblerait que les choix d'évaluation du bilan répondent davantage à un enjeu d'amélioration des informations fournies aux marchés de capitaux en Europe (et partant de leur fonctionnement : fluidité, homogénéité, réduction des asymétries d'information) que d'amélioration des informations fournies aux autorités de contrôle pour leur mission de protection des preneurs d'assurance.

2.2 Un provisionnement global conduisant à définir la faillite d'une façon qui a du sens pour l'actionnaire, non pour l'assuré

Cette opération de développement d'un cadre d'analyse transversal et de création d'un marché des portefeuilles d'assurance correspond à (et nécessite) une a-contextualisation des outils de mesure : des métriques indépendantes de leur finalité prudentielle, indépendantes des spécificités techniques du business

considéré¹¹³. Au-delà du *best estimate stricto sensu*, cette extraction du contexte est également à l'œuvre dans le passage du *best estimate* à la provision en valeur de marché.

Une entreprise en faillite est une entreprise qui n'est plus en mesure de rembourser ses créanciers donc, pour un assureur, d'indemniser ses assurés. Cette définition générale correspond à celle mise en avant par Solvabilité 2, la directive (EC, 2009, paragraphe 64) indiquant :

Le capital de solvabilité requis devrait être défini comme le capital économique que doivent détenir les entreprises d'assurance et de réassurance pour limiter la probabilité de ruine à un cas sur deux cent, ou alternativement, pour que lesdites entreprises demeurent en mesure, avec une probabilité d'au moins 99,5 %, d'honorer leurs engagements envers les preneurs et les bénéficiaires dans les douze mois qui suivent.

Est-ce réellement le cas ? Quelle est, dans les cadres Solvabilité 1 et Solvabilité 2, la définition d'une faillite et quelles conséquences cela emporte-t-il pour les assurés ?

La logique de détermination de la marge pour risque (MVM), qui, ajoutée au montant du *best estimate*, permet de fixer le niveau de provisionnement, a été débattue, oscillant entre deux options (ACAM, 2006, p57). La première option, testée dans la première étude quantitative d'impact (EIOPA, 2006), était de retenir l'écart entre le *best estimate* et un quantile donné des prestations actualisées (60%, 75 % ou 90%). Cette option aurait été centrée sur l'indemnisation des assurés : elle permettrait de garantir le fait que, en cas de faillite, c'est-à-dire après consommation du SCR, le montant global de provision serait tel que les assurés seraient indemnisés avec une probabilité donnée (60, 75 ou 90%). La seconde, celle qui a été choisie, était de considérer le coût d'immobilisation des fonds propres à détenir en regard des engagements. Cette solution finalement retenue est-elle pareillement porteuse

¹¹³ La financiarisation de la grille de lecture est telle que les concepts (actualisation de flux) pourraient être utilisés dans d'autres industries.

de sens pour les assurés, ou bien revêt-elle au contraire une signification concrète pour les actionnaires de l'entreprise ?

La MVM correspond à ce qui permet de passer d'un *best estimate* à une valeur d'échange. Rappelons que, comme il n'existe pas de marché de passif d'assurance¹¹⁴, cela est calculé à l'aide de modèles : on la détermine en actualisant les coûts annuels de capital à immobiliser, ces coûts correspondant au produit du montant à immobiliser¹¹⁵ et du taux de rémunération du capital (le CoC). Ce taux de rémunération du capital est un pourcentage forfaitaire, et le montant de la MVM lui est proportionnel. Certains représentants de l'industrie ont suggéré de retenir « 2 à 4 % » (EIOPA, 2008, p 111¹¹⁶) ; les compagnies européennes ont retenu en 2013 de 0,5 à 7% (Deloitte, 2014, p17) ; Allianz a changé année après année, retenant 5,5% en 2008, 6,5% en 2009 puis 4% à partir de 2010¹¹⁷ ; l'EIOPA a retenu 6% ; enfin, on considère de façon générale que les marchés financiers demandent une rémunération du

¹¹⁴ En matière de transfert de portefeuille, le transfert d'un engagement d'un assureur A à un autre assureur B ne peut être fait qu'après instruction d'un dossier, communication publique, possibilité d'opposition des assurés, et accord de l'autorité de contrôle. Ainsi, au-delà de l'aspect opérationnellement peu fréquent et financièrement non standardisé, les contraintes juridiques de protection des assurés ne permettent pas l'émergence d'un tel marché. De façon plus indirecte, si des opérations de titrisations sont apparues au cours de la dernière décennie, il s'agit à notre connaissance d'opérations structurées de gré à gré et sans marché secondaire ; qui plus est, l'assureur cédant des risques sur les marchés reste alors seul responsable vis-à-vis des assurés.

¹¹⁵ Il s'agit ici du principe, les détails apportant peu à la discussion. Deux déviations par rapport à ce principe peuvent cependant être notées : d'une part, on ne retient que les composantes non couvrables sur les marchés de l'exigence de capital, c'est-à-dire celles correspondant aux facteurs de risque de souscription et, d'autre part, on retient l'exigence de capital plutôt que le capital à immobiliser (qui correspondrait à l'exigence de capital nette de la VIF).

¹¹⁶ « A number of questions were raised regarding the appropriateness of the 6% cost of capital rate and the work of the CRO Forum was referenced by several undertakings. They argued that is questionable whether such a choice would lead to a reliable proxy for the cost of transferring a portfolio to a willing third party. Others felt that the cost-of-capital factor of 6% may overstate the true CoC for companies that may hold or acquire these liabilities, and argued for a factor in the range of 2%-4% instead. One undertaking thought that a single cost of capital rate which is the same across all lines of business and countries is unlikely to be appropriate.

Further consideration should be given to the appropriateness of the 6% cost of capital factor in light of the CRO Forum research.”

¹¹⁷ Allianz, 2008 à 2014, e.g. p50 en 2008, p 49 en 2009, p55 en 2010, p46 en 2014.

capital de l'ordre de 13 %. Il y a donc *a minima* un facteur 3 à 26 dans l'amplitude du CoC qu'on pourrait retenir, donc de la MVM qu'il pourrait être pertinent de déterminer.

Or, l'épaisseur de la MVM est ce qui sépare le montant des provisions, donc des actifs pour un organisme qui serait déclaré en faillite, du montant du *best estimate*, donc de l'espérance des engagements envers les assurés. La MVM a donc un impact direct sur la capacité à « honorer les engagements envers les preneurs est les bénéficiaires ». Selon son niveau, il restera plus ou moins d'actifs pour indemniser les assurés lorsqu'une entreprise sera déclarée en faillite. Ainsi, en fixant le montant du SCR, Solvabilité 2 ne fixe pas, contrairement à ses affirmations, la probabilité avec laquelle les assurés seront indemnisés : tout comme le niveau de provisionnement, ce montant ne signifie rien pour les assurés puisqu'à SCR donné, on peut avoir plusieurs niveaux de MVM différents, donc une protection plus ou moins importante des assurés. Contrairement à ce que signifie le considérant évoqué *supra*, *le SCR ne fixe pas la probabilité avec laquelle l'assuré est protégé* : celui-ci est plus ou moins bien protégé selon le niveau de CoC retenu dans la MVM alors que son niveau est arbitraire, indépendant du choix d'un niveau de protection des assurés à 99,5%.

Tautologiquement, la seule chose qu'on fait en fixant le montant du SCR, c'est fixer la probabilité avec laquelle les fonds propres économiques resteront positifs. La faillite pour Solvabilité 1 signifiait qu'il n'y avait plus assez d'argent pour indemniser les assurés. Dans Solvabilité 2, elle signifie qu'il n'y a plus suffisamment d'argent pour transférer le portefeuille, que, comme nous allons le voir, le seuil de l'option de vente détenu par l'actionnaire est franchi : l'objectif est devenu de comprendre dans quelles circonstances on peut transférer le portefeuille et non dans quelles circonstances on peut indemniser l'assuré.

Il serait aisé de présenter cela à l'aide d'un formalisme financier usuel, en recourant à des grandeurs très classiques pour la caractérisation d'option : l'action détenue par l'actionnaire est une option de vente (*put*) de seuil (*strike*)

nul dont le sous-jacent est la valeur des fonds propres économiques¹¹⁸ : tant que la valeur du sous-jacent reste positive, la valeur pour les actionnaires correspond à la valeur du sous-jacent ; lorsqu'elle devient négative, la responsabilité des actionnaires étant limitée, la valeur pour l'actionnaire devient nulle. En d'autres termes, les actionnaires détiennent une option de vente sur la valeur de l'entreprise de *strike* nul, et la probabilité fixée par le SCR est la probabilité de franchissement du *strike* : *ce que fait Solvabilité 2, c'est fixer la probabilité de franchissement par le sous-jacent du strike de l'option détenue par l'actionnaire.* Ici encore, on s'aperçoit que la notion de faillite et le niveau de sécurité fixé par Solvabilité 2 ne correspondent pas tant à des grandeurs relatives à la protection des assurés, qu'à des grandeurs relatives aux profils de *pay-off* de l'actif détenu par les actionnaires de l'organisme d'assurance. L'ambition de création d'un marché des passifs d'assurance semble avoir pris le pas sur l'ambition de protection des assurés.

2.3 Des fonds propres et exigences de capital associés à une notion de risque définie d'une façon qui a du sens pour l'actionnaire, non pour l'assuré

Le ratio de solvabilité est censé synthétiser la santé financière de l'entreprise. Ainsi, les conférences des autorités de contrôle sur les entités soumises à leur contrôle ont pour pierre angulaire de leur communication sur la santé financière des organismes les ratios de solvabilité des entreprises soumises à leur contrôle.¹¹⁹

Dans Solvabilité 1, la marge disponible correspond aux fonds propres sociaux : capital social, report à nouveau, etc. Les événements susceptibles de faire diminuer la marge disponible, donc d'altérer la santé financière de l'organisme d'assurance, sont donc ceux qui génèreraient un résultat négatif, c'est-à-dire des pertes. Dans Solvabilité 2 en revanche, par cohérence de construction entre les actifs, les provisions et les fonds propres, la marge disponible correspond aux

¹¹⁸ Dans le vocabulaire du CFO forum et des analystes financiers, il s'agit de la MCEV, qui est la somme de l'ANAV et de la VIF.

¹¹⁹ voir par exemple ACPR, 2014a, p42 à 47.

fonds propres économiques : elle contient la valeur du portefeuille (la VIF, correspondant aux perspectives de profits futurs nettes de la marge pour risque). Par conséquent, *les risques, c'est-à-dire les événements susceptibles de faire diminuer la marge disponible, ne correspondent plus aux pertes possibles, mais aux déviations possibles par rapport aux perspectives de profits futurs, c'est-à-dire à la crainte d'un manque-à-gagner*. Dans un tel schéma, quelle est la signification d'un élément tel que le ratio de solvabilité ?

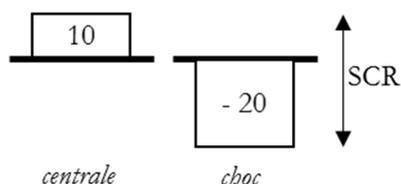
On ne peut plus désormais considérer que ce ratio soit un indicateur de la santé financière d'une entreprise. Prenons l'exemple de deux organismes exposés aux mêmes risques de pertes et dont la seule différence est un paramètre tel que l'un, ayant une perspective de profits futurs élevés, est davantage exposé au risque de manque à gagner ; voyons l'impact de ce paramètre sur, d'une part, la santé financière de chaque entreprise et, d'autre part, leur ratio de solvabilité.

Considérons deux entreprises en tous points identiques, hormis un : les contrats commercialisés par la compagnie A garantissent aux assurés de leur reverser 100% des bénéfices financiers, alors que les contrats de la compagnie B ne garantissent aux assurés que le reversement de 90% des bénéfices financiers. La compagnie A fait donc des promesses plus importantes que la compagnie B : il lui sera plus difficile de les tenir : elle est d'autant moins solide qu'elle a peu de perspectives de bénéfices, reversant tout aux assurés. Ainsi la compagnie A est dans une moins bonne santé financière que la B : elle est plus fragile. Quels vont être leurs ratios de solvabilité relatifs ? Revenons à la définition du risque sous Solvabilité 2 : il s'agit d'une déviation par rapport aux perspectives de profits futurs, d'un manque à gagner. Si les marchés financiers se portent bien, la compagnie A ne fait guère de profits puisqu'elle reverse tous les bénéfices financiers aux assurés, alors que la compagnie B fait d'abondants profits puisqu'elle conserve 10% des bénéfices financiers apportés par les actifs en face des provisions. Si les marchés financiers se dégradent, en cas de choc tel que simulé pour calculer les exigences de capital par exemple, il n'y aura pas de bénéfices financiers : les deux compagnies seront placées dans la même

situation et leur profits seront identiques. Ainsi, le risque, c'est-à-dire le manque à gagner en cas de choc, donc l'exigence de capital, sera bien plus élevée pour la compagnie A que pour la compagnie B. *La compagnie B (i) est plus solide et (ii) ne peut en aucune circonstance subir plus de pertes que A, mais elle aura une exigence de fonds propres plus élevée et peut avoir un ratio de solvabilité plus faible*, ainsi que le montre le schéma 2 ci-dessous.

Compagnie A

Capital social : 50
Taux de PB garanti : 100%



Situation centrale :

- Profits de 10 grâce aux frais de gestion

Situation de choc :

- Pertes de 20 pour maintenir les capitaux garantis

→ Montant de l'exigence de capital (SCR) :

Écart entre les deux situations (baisse des fonds propres économiques) : **30**

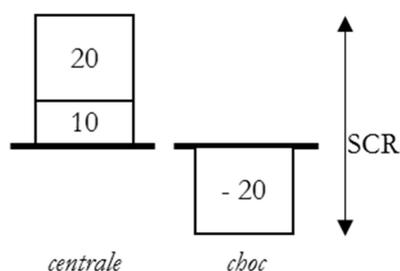
Marge disponible :

Capital social (50) + valeur en situation centrale (10) = **60**

→ **Ratio de solvabilité = 60/30 = 200%**

Compagnie B

Capital social : 50
Taux de PB garanti : 90%



Situation centrale :

- Profits de 10 grâce aux frais de gestion
- Profits de 20 grâce à la captation de 10% des bénéfices financiers

Situation de choc :

- Pertes de 20 pour maintenir les capitaux garantis

→ Montant de l'exigence de capital (SCR) :

Écart entre les deux situations (baisse des fonds propres économiques) : **50**

Marge disponible :

Capital social (50) + valeur en situation centrale (30) = **80**

→ **Ratio de solvabilité = 80/50 = 160%**

Schéma 7 :

illustration de la façon dont le ratio de Solvabilité peut diminuer si l'assureur renforce sa solidité financière en réduisant ses engagements

Ce mécanisme peut être observé dans de nombreux cas de figure. Ainsi, considérons de nouveau deux compagnies, en tous points identiques, hormis leur environnement fiscal. Contrairement à l'exemple précédent, nous nous intéressons ici à un paramètre exogène, qui n'est pas à la main de l'entreprise. Dans le pays A les sociétés sont soumises à un impôt sur les bénéfices important et dans le pays B, l'impôt sur les sociétés est faible. Dans un tel cas de figure, si un choc survient, la baisse des perspectives de bénéfices est plus faible dans le pays A. Ainsi, l'exigence de capital de l'organisme situé dans ce pays sera plus faible et le ratio de solvabilité peut également être plus avantageux¹²⁰.

Une fois de plus, on mesure combien l'information fournie par le ratio de solvabilité est peu en adéquation avec un enjeu d'exposition de l'assuré au risque de non indemnisation ou non remboursement. Ce ratio de solvabilité devient orienté vers la captation et la mise en regard de la valeur de l'entreprise et de ses variations. *Ainsi, l'information synthétisée par le ratio de solvabilité se rapproche davantage sous Solvabilité 2 d'un ratio de Sharpe¹²¹ informant les investisseurs sur le couple risque/rendement du titre qu'est l'action d'une société d'assurance* : un indicateur synthétisant les perspectives de gains futurs, ajustées de leur volatilité.

¹²⁰ Dans les cas de figure présentés (on pourrait prendre d'autres exemples, par exemple la commercialisation d'unités de compte, où les assureurs ne portent pas de risque de placement, peut être désavantageuse car la plus grande volatilité des encours peut générer une plus forte volatilité des chargements de gestion prélevés, donc des profits), l'exigence de capital évoluera toujours dans le « mauvais » sens. Le signe de l'évolution du ratio de solvabilité, lui, dépendra des variations relatives de l'exigence de capital et du capital disponible et sera donc fonction de l'interaction des différents paramètres techniques et financiers, sans qu'il soit possible de dégager d'orientation systématique.

¹²¹ Le ratio de Sharpe est un indicateur répandu de mesure de la performance ajustée du risque d'un portefeuille d'actif, et correspond au rapport entre le sur-rendement du portefeuille par rapport au taux sans risque divisé par sa volatilité.

2.4 Quand une crise survient...

L'analyse que nous venons d'établir pourrait être critiquée du fait qu'elle procède par une analyse du bilan poste à poste, détachée d'une vision d'ensemble du nouveau système bilanciel. Il manquerait donc à ce stade le bouclage qui donne toute sa pertinence au système *en tant qu'outil de protection des assurés*. De fait, un rappel de la vision sous-jacente à cette approche pourrait faire davantage apparaître la structure comme étant pensée pour l'assuré : en s'attachant à la valeur de transfert de chaque bloc du bilan, on met sous contrôle le fait que, si une compagnie a une marge disponible inférieure au minimum requis (et qu'elle doit donc cesser son activité), alors il devrait être possible, tant que cette marge est encore positive, de transférer son portefeuille à un ou plusieurs autres assureurs en leur transférant tout à la fois ses passifs et les actifs en face. Ainsi, l'idée sous-jacente est que les engagements pouvant être repris, d'un bloc ou portefeuille par portefeuille, les assurés sont bien protégés. Dans un tel schéma global, l'évaluation des postes de passif et d'actif serait bien un moyen au service de la fin qu'est la protection des assurés.

Pour qu'une telle réponse soit valable, il faudrait (et il suffirait) que la valeur du passif soit *effectivement* un prix de transfert : que la mise en regard des passifs et des actifs permette *effectivement* le transfert du portefeuille.

Toutefois, Schleiffer et Vishni (1992) ont montré que, lorsqu'une entreprise financière était en détresse, ses pairs tendaient à l'être également, ce qui rendait non pertinent le fait de considérer la valeur de marché d'un collatéral et plaidait pour la valorisation des collatéraux à leur valeur de liquidation. Certes, le principe du SCR dans Solvabilité 2 est précisément de s'assurer qu'*après* un choc d'amplitude donnée, les fonds propres ne seraient pas négatifs, ce qui assurerait la transférabilité des portefeuilles. Mais ceci passe à côté du point central de la critique de Schleiffer et Vishni : le fait qu'un actif ne peut être cédé que si un acheteur est disposé à l'acquérir, et le fait que la santé de l'acheteur et sa disposition à payer, donc le prix de transfert, sont fonctions non seulement

des caractéristiques de l'actifs mais aussi, et de façon cruciale, du contexte (e.g. période de crise ou non).

En régime de croisière, déjà, il n'existe pas de marché des passifs d'assurance : si une entreprise cède l'un de ses portefeuilles à une autre, ce qui est très rare et nécessite l'accord de l'autorité de contrôle, il s'agit toujours d'une transaction de gré à gré, et avec mobilisation de nombreux conseils analysant les contrats, les modalités de provisionnement, etc. C'est la raison pour laquelle la valeur de marché des provisions et des fonds propres est construite à l'aide de modèles et des conventions associées, et non constatée comme la valeur des placements. En cas de choc, par exemple de catastrophe naturelle, de choc action ou immobilier, de baisse des taux ou de hausse des spreads, la valeur des passifs sera recalculée avec les mêmes outils de projections de *cash-flows* : la valeur de marché des passifs est déterminée à partir de modèles qui, étant une actualisation de flux, par construction ne font en rien intervenir l'existence d'acheteurs potentiels. En d'autres termes, dans la valorisation des passifs destinée à mesurer leur valeur dite de marché en situation actuelle (mesure des fonds propres disponibles) et en cas de stress (mesure de l'exigence de fonds propres), la baisse du nombre d'acheteurs n'intervient en rien. Elle ne peut être captée : on a créé une valeur dite « de marché » indépendante de l'existence et du nombre de repreneurs éventuels. Comment interpréter cela ?

En constatant tout d'abord que, de ce fait, la valeur de marché est ici virtuelle. Cette grandeur n'est pas opérationnelle en vue d'un transfert effectif de portefeuille : puisque le prix a été déterminé indépendamment de l'existence (voire en niant l'absence) d'acheteur¹²², le fait qu'en cas de choc, lorsque toutes

¹²² Et suivant, cela est inévitable dans un contexte réglementaire, une batterie de conventions qui n'ont aucune raison (i) de correspondre à celles d'un acquéreur, en termes de taux de rendement ni (ii) de « frontière des contrats » par exemple, ou encore d'être réalistes, par exemple lorsqu'on considère que le taux de rachat ne dépend que du taux servi, même en cas de scénario de crise où les assurés auraient donc toutes les raisons de paniquer.

les entreprises d'assurance seront en difficulté¹²³, le prix calculé et les actifs en représentation permettront *effectivement* de fonder un transfert n'est pas crédible. Il s'agit donc d'une simple *mesure* mais qui, virtuelle, ne peut fonder une *action* (si transfert il y avait, il n'aurait aucune raison de se faire suivant cette valorisation). Or la possibilité *effective* de mise en œuvre d'une action (en l'occurrence, de transfert de portefeuille) est *la condition de pertinence pour la protection des assurés* de cette métrique.

Ceci brise donc la logique du bouclage de Solvabilité 2 comme assurant une protection des assurés *via* une valorisation des provisions en valeur de marché *par la possibilité de transfert*, puisqu'on ne peut mettre concrètement cette valorisation des engagements en regard de celle des actifs qui les représentent pour permettre leur reprise par un tiers. En d'autres termes, l'information fournie par la valorisation des passifs en valeur de marché correspond non à une estimation *absolue* d'un prix de transfert, mais à une évaluation *relative* des différents portefeuilles, un benchmark pour l'investisseur.

Le passage de Solvabilité 1 à Solvabilité 2 marque donc l'abandon d'une valorisation des provisions fondées sur l'engagement vis-à-vis des assurés pour un pseudo prix de transfert qui n'a de valeur que relative : cette valeur ne saurait être mise effectivement en regard des actifs pour permettre la reprise des engagements par un tiers. Ainsi, lorsqu'une crise survient, si en Solvabilité 1 la marge absorbe le choc de telle sorte qu'il reste suffisamment d'actifs pour indemniser les assurés, en Solvabilité 2 en revanche, la marge absorbe les chocs jusqu'à permettre, non une indemnisation des assurés, mais un transfert virtuel, sans marché d'échange. *De facto*, ce que permet le système quantitatif de Solvabilité 2 dans son ensemble n'est donc pas tant d'assurer une transférabilité des portefeuilles d'assurance pour protéger les assurés qu'une possibilité de comparaison de flux de *cash-flows* futurs suivant un jeu donné de conventions.

¹²³ Cela peut notamment s'observer en cas de krach lent, par exemple actuellement sur les portefeuilles retraite avec l'affaissement long et continu des taux.

3. Quelques conséquences possibles

Cette évolution de l'appréhension du bilan d'un assureur et le changement de vision de l'activité d'assurance qu'elle véhicule modifient les regards extérieurs (des analystes, des régulateurs, etc.), et pourraient donc tendre à faire évoluer le pilotage des compagnies d'assurance, soit en déplaçant les contraintes, soit, de façon plus fondamentale, en modifiant la fonction objectif du *top management*. Qu'en est-il ?

Solvabilité 2, catalyseur pour l'assurance d'un processus économique de fond

Une responsable d'équipe au sein des directions du pilotage d'une filiale d'un grand groupe d'assurance explique :

Avec Solvabilité 1, à [sa compagnie], on cherchait les profits, mais sans regarder l'exigence de capital. Là, on a appris à regarder la rentabilité.

Une telle évolution de la mesure de la performance n'est pas spécifique à l'assurance et avait débuté bien avant Solvabilité 2. Ainsi, vantant Solvabilité 2 comme « une opportunité de pilotage de la performance », Foulquier (2009) écrit :

Cette étude concerne tous les acteurs du secteur de l'assurance, quelles que soient leurs caractéristiques, puisque finalement, ils sont tous contraints à renseigner les nouvelles normes prudentielles de Solvabilité II et ont tous comme objectif la satisfaction des actionnaires ou des sociétaires.

Depuis la révolution industrielle, la mesure de la performance des entreprises a fait l'objet de vastes travaux. La création de valeur est devenue l'une des pierres angulaires des sciences économiques dès le début du XIX^{ème} siècle et s'est imposée progressivement dans le monde de l'entreprise dans la seconde moitié du XIX^{ème} siècle. Contrairement aux indicateurs classiques de mesure de valeur basés sur le chiffre d'affaires, le résultat net ou les marges opérationnelles, la création de valeur est évaluée en considérant que les ressources des entreprises (capitaux propres et / ou dettes) ont un coût. Cette approche s'applique à l'ensemble des entreprises de tous les secteurs.

A titre illustratif, certains assureurs considèrent que l'activité de responsabilité civile générale génère une performance satisfaisante parce qu'elle dégagne une marge nette (résultat net / chiffre d'affaires) parfois jusqu'à trois fois supérieure à celle de l'assurance dommages automobile des particuliers. Pourtant, lorsque cette marge nette est mise au regard des capitaux économiques alloués déterminés à partir de la volatilité de l'activité, l'assurance dommages automobile des particuliers apparaît souvent plus rentable. Ainsi, considérer le coût du capital dans la mesure de la performance est de nature à imposer aux cadres dirigeants mais aussi aux opérationnels, une certaine discipline capitaliste tout au long du cycle d'exploitation, car l'utilisation de ce capital n'est pas gratuite.

De fait, Solvabilité 2 semble agir comme un catalyseur puissant pour intégrer une telle approche dans les modes de réflexion des dirigeants des compagnies d'assurance. Ainsi, le P-DG de la Scor déclare :

« Le Groupe sera géré sur le principe de la maximisation de la valeur économique. [...] »

« S2 pousse les entreprises à mieux gérer leur capital, c'est pour ça qu'en tant qu'économiste, j'aime bien ça. Dans S1, on était au-dessus ou en dessous. Dans S2, c'est à vous de définir le niveau de capital que vous allez devoir gérer. C'est pour ça qu'il y a en Europe cette obsession de la gestion du capital, mais pas pour le régulateur, pour la société elle-même ! Donc le capital doit devenir la pierre angulaire de la construction. On utilise le capital comme outil de pilotage. L'acquisition des sociétés X et Y aux Etats-Unis a été faite avec le modèle interne. Et c'est le modèle interne qui nous a permis de calibrer les conséquences avec toutes les interactions avec toutes les régions du Groupe. [...] C'est un formidable outil stratégique. [...] Maintenant, les plans de souscription passent par le modèle interne »

(Kessler, D. 14 octobre 2014, Journée *Scor Global Life - Solvabilité 2*)

Ainsi, Solvabilité 2 semble constituer un outil qui « modernise » le pilotage des assurances en y introduisant un souci particulier des fonds propres immobilisés et une mesure de la performance centrée sur la rentabilité sur ces fonds propres. Suivant quelles modalités, avec quels impacts ?

Une logique de disparition des dirigeants ?

Cette modernisation des critères de pilotage ne peut être dissociée d'une logique d'évolution du rapport de force entre dirigeants et actionnaires, évolution qui passe par deux canaux : la maîtrise de l'information, et le choix du cadre de pensée.

En non vie, le passage de provisions « prudentes », dont le niveau de suffisance était laissé à l'appréciation des dirigeants, à un montant de provision fondé sur un *best estimate* corrigé d'une MVM semble marquer une perte de maîtrise du pilotage de l'entreprise par les dirigeants : une diminution de la capacité à gérer les risques en conservant du « gras » dans les provisions les années fastes pour absorber les conséquences des années difficiles et l'atrophie d'un outil fondamental de pilotage du résultat. Les dirigeants deviendraient ainsi comme transparents, les analystes ayant directement accès à l'information fournie par les experts.

Si, en vie, le caractère moins mécanique du provisionnement sous Solvabilité 2 par rapport à Solvabilité 1 semble jouer en sens inverse, l'accroissement considérable de complexité (tel qu'aucun maillon de la chaîne allant de l'expert au CEO ne peut désormais s'approprier les résultats) pourrait s'interpréter comme un alignement de la connaissance de l'activité qu'avaient les dirigeants sur la méconnaissance qu'en avaient les actionnaires. Surtout, en vie, les provisions Solvabilité 2 sont déterminées sur la base de flux projetés sous probabilité risque neutre et non sous probabilité historique : ceci a deux conséquences :

- faire obstacle à toute interprétation concrète des résultats, ce qui rejoint le point *supra*,

- et être conforme aux pratiques de marché pour le *pricing* d’option mais introduire une incohérence insurmontable dans les interactions actif/passif¹²⁴.

On voit ici émerger le second point (le fait que l’évolution du rapport de force passe également par l’alignement du cadre de pensée des dirigeants sur ceux des analystes quantitatifs) : Solvabilité 2 aligne les modalités de calcul des exigences de capital sur le cadre cognitif développé par les mathématiques financières et utilisé sur les marchés financiers.

Ainsi, un des principes de base de l’assurance qui repose sur la loi des grands nombres est que les portefeuilles sont d’autant plus stables qu’ils sont de taille importante. Ce phénomène de stabilisation relative du résultat en cas de croissance de la mutualisation était capté par Solvabilité 1¹²⁵. *A contrario*, l’EIOPA (2013) recommande que les mesures de risques retenues en cas de modèle interne soient homogènes¹²⁶, c’est-à-dire que, par exemple, le doublement de la taille d’un portefeuille conduise à une mesure de risque deux fois plus élevée. Ceci est contraire aux principes que nous venons d’évoquer, *mais cohérent avec les outils des mathématiques financières* où, lorsqu’un investisseur envisage d’investir un montant sur un actif, le couple risque/rendement de celui-ci est indépendant du montant investi.

¹²⁴ Lorsqu’on *price* une option, il n’y a pas d’interaction actif/passif, mais uniquement des cash-flows d’actifs. Lorsque cela est transposé à l’assurance, les actifs sont par mimétisme risque-neutralisés et rapportent le taux sans risque, mais les modèles les font alors interagir avec des passifs qui, eux, ne peuvent pas (et donc n’ont pas) été risque-neutralisés : on ne remplace pas, par exemple, les taux garantis par le taux sans risque.

¹²⁵ Voir supra 1.1.5 et CE (73), art. 16, §3.

¹²⁶ “[N]ational competent authorities can form a view on how the undertaking describes the risk metric in respect of the following risk measurement properties: [...]”

- Homogeneity: multiplying the size of a portfolio by a scalar x the risk measure is multiplied by x ; [...]

Without requiring the risk metric to follow the properties above, a detailed description of circumstances where the risk metric would not follow one or more of them could be asked by national competent authorities to form a view on the appropriateness of the approach followed by the undertaking”, EIOPA (2013), p 97, § 5.213.

De même, l'agrégation des différents risques, donc, pour les risques de souscription non vie, des différentes lignes d'activité, est réalisée par des matrices de corrélation¹²⁷, là encore de façon parfaitement alignée avec les outils usuels de mathématiques financières issus du CAPM et alors que cette méthodologie d'agrégation ne permet pas, de façon générale, d'agréger des VaR. En d'autres termes, Solvabilité 2 sanctuarise une diversification telle que modélisée dans la théorie financière classique. Ceci devrait avoir pour conséquence que, en optimisant leur allocation de capital entre différents portefeuilles de passifs et la rentabilité sur ce capital global, les entreprises d'assurance opèrent une diversification *répondant aux critères des analystes financiers*. Ce à quoi devrait conduire la logique de Solvabilité 2, tant dans l'appréciation des passifs et profits futurs par une espérance de flux futurs actualisés en risque neutre que par ses choix de quantification du risque et des modalités de diversification, c'est donc de *faire internaliser au sein des organismes d'assurance et au niveau des prises de décision les concepts utilisés par les analystes tarifant les options financières et faisant une gestion de portefeuille quantitative*.

Ceci tendrait à faire de la direction générale, et notamment du directeur financier, non plus un stratège industriel et commercial agençant stratégiquement les lignes de produits, mais un gestionnaire d'actifs optimisant le couple risque/rendement global d'un portefeuille de titres (les lignes d'activité) ayant chacun son couple risque/rendement déterminé : l'entreprise d'assurance deviendrait un véhicule de titrisation faisant de l'intermédiation de couple risque/rendement, un fonds de risques ayant une volatilité et un rendement global.

Ceci est-il de nature à améliorer la prise en compte du risque par les dirigeants, conformément à l'ambition initiale de Solvabilité 2 ?

¹²⁷ E.g. EC, 2010, QIS5 Technical Specifications, p95, 108, 148, 168, etc.

Vers une disparition de la prise en compte du risque ?

Sous Solvabilité 1, le risque n'est pas pris en compte dans le pilotage *via* un impact au dénominateur du ratio de rentabilité, puisque l'exigence de capital n'est pas conçue pour le refléter. Le risque ne peut donc pas être pris en compte explicitement sous la forme d'une mesure de rentabilité *ex ante* ajustée du risque, mais uniquement par la crainte que la perspective de sa réalisation fait porter sur le résultat (numérateur du ratio de rentabilité).

D'un point de vue qualitatif, il semble que les organismes d'assurances prenaient en compte la perspective d'événements adverses dans leurs décisions. De fait, on constate que les risques identifiés comme prépondérants par Solvabilité 2 étaient déjà sous surveillance. En vie, il s'agit des risques de marché. Or, à l'actif, les compagnies ne saturaient pas les limites de dispersion réglementaires¹²⁸, faisant preuve d'une prudence supérieure à la prudence minimale imposée par la réglementation. En non vie, il s'agit des risques de souscription : catastrophe, prime et provision. Pour gérer les risques de catastrophes, elles bâtissaient un programme de réassurance (bien que ceux-ci fussent structurellement coûteux pour les assureurs puisque rentables pour les réassureurs) ; le risque de provision était géré par le « gras » susceptible d'être dérogé en cas d'événement adverse ; quant au risque de prime, il était géré à la fois par les deux mécanismes comptables mentionnés *supra* et par le fait que, en cas de lancement d'un nouveau produit, la montée en puissance du chiffre d'affaires était progressive afin d'affiner progressivement le tarif au fil de l'acquisition d'information sur la rentabilité. Plus globalement, les organismes immobilisaient bien plus de fonds propres que la réglementation ne l'exigeait, ce qui renforçait encore leur résistance aux chocs. Ainsi, un ancien contrôleur explique :

¹²⁸ Considérons par exemple la poche action, réputée avoir le meilleur rendement moyen. L'exposition moyenne aux actions du marché français était à fin 2006 inférieure à 30 % (ACAM, 2006, p20) alors que la limite de dispersion par catégories était de 65% des engagements réglementés.

Au début des années 2000, nous [les contrôleurs européens] avons partagé des informations sur les bonnes pratiques de nos industries. Ils [nos homologues] voulaient comprendre comment nos entreprises [les entreprises françaises] avaient résisté à Lothar et Martin [l'épisode de tempêtes exceptionnel ayant notamment balayé l'Ouest et le Nord de la France fin 1999]. Et effectivement, certaines ont perdu trois ou quatre fois l'exigence de capital. Mais c'était très simple : elles avaient six ou sept fois l'exigence.

A défaut d'être systématiquement explicitée de façon formelle, la gestion des risques pouvait alors être intégrée à l'activité de l'entreprise : si nous reprenons la liste évoquée ci-dessus, en lien avec la direction générale, les équipes d'investissements dispersaient leurs actifs, les équipes responsables du provisionnement prenaient des marges de sécurité, les équipes lançant un produit veillaient à la connaissance qu'ils avaient du risque et à la souscription de traités de réassurance, etc.

Solvabilité 2, dans la continuité de la logique d'alignement des modes de pensée des dirigeants sur ceux des actionnaires que nous avons présentée, tend à promouvoir, ne serait-ce que par le poids relatif du pilier quantitatif dans les préoccupations des dirigeants, une centralisation de la gestion des risques dans les fonctions de pilotage. Il s'agit désormais d'une activité relevant de la holding, du pilotage financier. Considérons pour illustrer cela les fonctions et parcours des directeurs des risques des principaux groupes d'assurance européens :

- Les Group CRO d'Axa (de Mailly-Nesle) et Generali (Panizza) étaient auparavant les CFO de filiales (respectivement de la zone Méditerranée/Amérique Latine, et d'InassItalia),
- A Allianz, il n'y a pas de CRO au *Board of Management*, mais un CFO ayant la responsabilité à la fois des finances et des risques (Wemmer),
- A Aviva, le dernier CRO (Lister) était auparavant CFO d'une filiale puis est devenu CFO d'une autre filiale. Lorsqu'il occupait les fonctions

de CRO, sa fonction exacte était « Chief Risk *and Capital* Officer » (je souligne).

De fait, dans le système Solvabilité 2, la gestion des risques devient équivalente à la gestion de capital. Cette quantification du risque par une mesure homogène agrégée ne constitue pas un *early warning* permettant d'identifier rapidement une dérive pour la corriger, mais répond à la vision d'une holding souhaitant comparer des investissements et rejoint en ce sens les observations de Power (2009) réalisées dans un champ autre que l'assurance. Solvabilité 2 tend à promouvoir une gestion des risques alignée sur l'acceptation qu'en ont les mathématiques financières : la fonction du CRO n'est plus de *prévenir* les risques liés à l'activité en identifiant des dangers donnés et en entreprenant des actions associées, mais de constater, de *mesurer* des risques, qu'on considère comme statistiquement inéluctables, pour contribuer à l'optimisation du couple risque/rendement de l'actif dans lequel investissent les actionnaires.

Cette dimension a été internalisée par certains opérationnels. On peut citer, à titre d'exemple, cette réponse d'un représentant d'une direction métier à un représentant de la direction des risques qui souhaitait la modification d'un contrat qu'il jugeait trop risqué pour l'entreprise :

« Mais vous modélisez les risques. Vous avez fait vos calculs et c'est dans l'exigence de capital. Donc puisqu'il y a du capital en face, le risque est déjà géré, alors arrêtez de nous embêter avec. »

Qu'elle soit rhétorique ou non, cette réponse tend ainsi à montrer que, paradoxalement, l'instauration d'une quantification précise des risques et d'une exigence de capital associée peut conduire à dégrader la gestion opérationnelle des risques.

L'état d'esprit des régulateurs est également modifié

Il est par ailleurs frappant de constater à quel point le schéma de pensée des régulateurs a évolué entre le début des discussions sur Solvabilité 2 et son entrée en vigueur. Un symptôme, à défaut d'une preuve, en est l'évolution de leur position sur l'arbitrage entre absence de biais (justesse) et biais favorable

aux assurés (protection). Nous avons mentionné (1.2) des discussions entre régulateurs sur l'utilité ou non d'avoir une vision juste, au tournant des années 2000 et au tournant des années 2010. L'équilibre semble désormais s'être déplacé, avec le reproche de choix prudents formulés à des organismes d'assurance par des équipes des autorités de contrôle. Ainsi, en 2014, une observation participante a conduit à noter les deux entretiens suivants, dans le cadre d'une procédure d'approbation d'un modèle interne par des autorités de contrôle, la première française, la seconde allemande, qui montrent une acculturation des équipes des autorités de contrôle et la mesure de leur appropriation de ces schémas de pensée :

- « Votre hypothèse n'est pas réaliste, vous devriez l'affiner.
- Oui mais ça nous pénalise [en sous-estimant la marge disponible ou bien en calculant une exigence de capital plus élevée qu'il ne serait nécessaire]. On préfère avoir quelque chose de simple, qu'on comprend bien, et en choisissant cette hypothèse-là, on est conservateurs [i.e. prudents].
- Oui, mais ce n'est pas ce que dit la réglementation. Il faut que ça reflète la réalité. »

- « Pourquoi vous faites cette hypothèse-là ?
- Elle est prudente.
- On ne vous demande pas d'être prudents, on vous demande d'être justes. »

Conclusion

Synthèse

Solvabilité 2 met en place une grille d'analyse ne visant pas tant à assurer l'indemnisation des assurés qu'à informer les investisseurs en les plaçant sur un pied d'égalité :

- Le *best estimate*, cœur des provisions, n'est plus, en vie, ce que l'assureur doit impérativement garantir à chaque assuré, mais devient la mesure

permettant de déterminer comment les actifs seront répartis entre la communauté des assurés et les actionnaires.

- La faillite ne représente plus le seuil où l'assureur ne sera plus en mesure de tenir ses engagements vis-à-vis de ses créanciers, au premier rang desquels les assurés, mais le *strike* de l'option de vente sur les fonds propres économiques qu'est l'action détenue par les investisseurs.
- Le risque n'est plus celui de perdre de l'argent, mais celui de réaliser moins de bénéfices qu'anticipé : il s'agit d'une vision financière du risque en tant que volatilité et non d'une vision industrielle du risque en tant que danger. Le couple {fonds propres économiques ; exigence de capital} s'apparente à un couple rendement/risque. Le ratio de solvabilité devient proche d'un ratio de Sharpe renseignant l'investisseur sur l'équilibre entre ces deux grandeurs. On peut par exemple le faire diminuer tout en rendant la santé financière meilleure dans tous les états futurs du monde possibles, par exemple lorsqu'on réduit les engagements ou les impôts (et que les profits ne peuvent donc être que plus élevés), parce que la volatilité de cette santé toujours meilleure est proportionnellement plus importante.

L'analyse de cette étude s'est concentrée sur le Pilier 1 de Solvabilité 2, c'est-à-dire la quantification du bilan et des exigences de capital associées. Le pilier 3, centré sur les reportings devant être produits par l'assureur et qui introduit, par rapport à la réglementation précédente, l'obligation de communiquer des batteries d'informations quantitatives au public, peut difficilement ne pas être rapproché de cette ambition d'informations aux analystes. Le discours permettant de l'articuler à une visée de protection des assurés est celui de la « discipline de marché » : se sachant observé par ses pairs, l'assureur limiterait sa prise de risque. Il s'agit là d'une transposition de pratiques issues de l'univers bancaire. Dans ce domaine originel, les ressources d'une banque viennent des autres banques : le besoin de refinancement de chacune génère des interactions

telles qu'effectivement, chaque banque a intérêt à analyser la solidité financière de ses consœurs avant toute opération et qu'une banque perçue par les autres comme dangereuse verra augmenter le coût de ses ressources. On comprend donc bien dans quelle mesure, dans le secteur bancaire, la communication d'information aux marchés peut engendrer une auto-discipline conduisant chacun à limiter ses risques pour limiter ses coûts. Dans le domaine de l'assurance, il n'y a pas de telles relations entre les organismes. Un assureur mal évalué par ses pairs ou par les analystes ne verra pas le coût de ses ressources quotidiennes augmenter. *A contrario*, dans le domaine de l'épargne particulièrement, diffuser des informations alarmistes aux marchés est susceptible de conduire à l'équivalent de *bank run*, incitant les déposants à retirer leurs encours et introduisant dans l'activité d'assurance des risques qui auparavant n'existaient que dans le secteur bancaire¹²⁹. Ainsi, du point de vue de l'assuré, la non-transparence est une sorte de bien public, et les assurés ont, collectivement, intérêt à ce que seule l'autorité de contrôle ait accès aux informations permettant d'apprécier la santé financière des organismes. L'argument de la discipline de marché, dans le cadre de Solvabilité 2, ne peut donc pas être recevable du point de vue de la protection de l'assuré : le pilier 3 également semble être aligné sur l'intérêt de l'investisseur et non celui de la communauté des assurés.

Par ailleurs, nous avons vu qu'au-delà et du fait de ces caractéristiques techniques, le pilier 1 de Solvabilité 2, par ses impacts putatifs, pourrait marquer la transformation d'une activité industrielle vers une activité financière, dans la façon dont elle est menée, dans la façon dont elle est pensée et arbitrée par ses dirigeants. Elle véhicule un mode de pensée incitant les dirigeants des entreprises d'assurance à se comporter en gestionnaire de

¹²⁹ Ceci peut générer une faillite, par exemple si, lors d'une remontée brutale des taux, un assureur bascule en moins-values latentes obligataires et qu'il doit réaliser ces moins-values alors que la nature de l'assurance, de long-terme, aurait pu lui permettre de détenir ces obligations à maturité.

portefeuille de passifs d'assurance, à interioriser les modalités d'analyse et faire percoler les critères de décision de la finance quantitative pour les répliquer à leur niveau. Ceci consisterait notamment à prendre acte des risques et à tenter d'optimiser le rendement associé en les diffusant, non à les réduire pour les mettre sous contrôle. Ceci est conforme avec le discours tenu par le numéro deux d'un organisme d'assurance important : « je gère une grosse SICAV » [2011].

Comprendre pourquoi

Au début des années 2000, les pratiques de marché constituaient un terreau favorable à la mise en œuvre de tels choix. A cette époque, la réassurance *finite* commençait à être suffisamment développée pour générer des prises de position des autorités de contrôle. Son objectif était d'embellir les bilans et résultats des assureurs aux yeux des marchés financiers en matérialisant dans les comptes des perspectives de profits futurs. Solvabilité 2, en valorisant les fonds propres par des méthodes financières, ne fait alors que systématiser ces pratiques en alignant la grille de lecture réglementaire sur la grille de lecture financière¹³⁰. Mais il sera intéressant de parvenir à comprendre, comment, à partir d'un terreau favorable, dont l'aspect plus large, cognitif et institutionnel, a été décrit par François (2015), ces options ont été retenues.

Certes, lors de la conception de Solvabilité 2, les régulateurs ont fait le choix délibéré, contrairement aux régulateurs américains, de ne pas imposer un double jeu de comptes aux assureurs, l'un prudentiel, l'autre pour le reporting financier (François, op.cit.). Ceci pourrait expliquer un tel alignement des choix comptables sur des concepts ayant du sens pour les investisseurs : les

¹³⁰ Dans le temps long, les grandes évolutions comptables s'inscrivent dans une tendance d'accélération de la distribution des profits telle que décrite par Richard (2015). Initialement, la valeur retenue était celle de la valeur à la casse : les remontées de dividendes n'étaient possibles qu'après l'absorption des investissements initiaux. Puis, la valorisation choisie a été le coût historique : le coût des investissements initiaux était amorti, lissé dans le temps, permettant une matérialisation des profits plus rapides. Enfin, la juste valeur permet de considérer l'espérance de profits futurs comme un profit : elle alimente les fonds propres et devient distribuable avant même qu'ils ne soient réalisés.

régulateurs de l'assurance auraient pris la seule option à leur disposition pour construire tant bien que mal sur cette base un système prudentiel. Toutefois, la chronologie relative des évolutions de la comptabilité prudentielle et de la comptabilité financière n'est pas compatible avec une telle hypothèse : La norme IFRS 4, publiée en 2004, n'a fait basculer en valeur de marché que les actifs des sociétés d'assurance. Dès le QIS 1 (en 2005), Solvabilité 2 avait acté le principe de provisions en *best estimate* et dès fin 2006, il apparaissait que la MVM serait déterminée en fonction du coût en capital, entérinant le principe d'une valorisation globale en valeur de marché (ACAM, 2006), à l'instar du *Swiss Solvency Test* lancé en 2005. Or ce n'est qu'en 2010 que le régulateur comptable a lancé le chantier des normes IFRS 4 phase 2 visant à valoriser les passifs d'assurance sur la base d'un *best estimate*. Et si, fin 2015, cette ambition semble en passe d'être atteinte, c'est notamment par un retournement de l'argumentation : le caractère souhaitable d'un alignement des normes financières sur la norme prudentielle existante, Solvabilité 2. Il apparaît donc que les régulateurs de l'assurance n'ont pas subi ces choix : ils ont été proactifs dans le développement d'une vision pertinente pour l'actionnaire plutôt que pour l'assuré.

Alors que les considérants de Solvabilité indiquaient très explicitement que « *la stabilité financière et la stabilité et l'équité des marchés* » constituaient « *d'autres objectifs [...] qui devraient être pris en compte sans détourner cependant du principal objectif* » il est désormais patent que, dans la mise en œuvre de ses concepts si ce n'est dans sa philosophie, les priorités entre la mission cœur et la seconde mission accessoire ont été inversées. Là où des régulateurs comptables eux-mêmes jugent que

« L'information comptable n'est pas un outil de régulation du marché, c'est aux régulateurs d'appeler à la prudence. »¹³¹,

¹³¹ P. Danjou, membre du Board de l'IASB, cité par O. Dufour (2013) dans Normes IFRS – L'IASB tente de répondre aux critiques, Option Finance n°1225 – 10 juin 2013.

chacune des composantes quantitative de Solvabilité 2 a été pensée pour l'actionnaire, le renseigne, mais ne correspond pas à l'option qui aurait été retenue si le choix avait été orienté par un objectif central de protection des assurés. La question reste donc entière : comment expliquer que, dans la pratique si ce n'est dans l'ambition, l'objectif principal de Solvabilité 2 n'ait pas été de garantir la protection des bénéficiaires et que cette mission ait été sacrifiée au profit de l'objectif initialement secondaire d'information aux marchés¹³².

Deux hypothèses, non exclusives l'une de l'autre, peuvent être avancées. L'une est celle d'un choix délibéré résultant d'une visualisation de la protection des assurés comme devant être mise en œuvre avec voire par la construction d'un marché fluide des passifs¹³³. Cette approche serait crédibilisée d'une part, par mimétisme avec l'état d'esprit ayant présidé à la conception de la réglementation Bâle 2 dans le domaine bancaire à la fin des années 90 (Baud, 2013 ; Baud et Chiapello, 2015)¹³⁴ et, d'autre part, par la cohérence théorique d'une architecture globale où dès lors qu'on s'assure de la possibilité d'une reprise des engagements, on s'assure du fait que les engagements seront honorés. Cette matrice de pensée peut ainsi être suffisamment convaincante pour que, au-delà de constituer un simple terreau favorable, il soit assumé, en

¹³² S'il est permis de douter que l'objectif de stabilité financière, cité comme première mission accessoire, sera atteint au vu de la procyclicité des exigences quantitatives de Solvabilité 2, tel n'est pas l'objet des présents travaux. La seconde mission annexe, « la stabilité et l'équité des marchés », est ambiguë, pouvant faire référence soit aux marchés des produits d'assurance, soit aux marchés financiers. Nous l'interprétons dans la seconde acception. Si elle faisait référence à la première, notre étude signifierait que Solvabilité 2, en se souciant de l'information aux marchés, a été détournée de son objectif principal à l'occasion de la poursuite d'une fin qui n'était ni dans ses objectifs principaux, ni dans ses objectifs accessoires.

¹³³ Objectif qui, en soi, indépendamment de la protection des assurés, serait cohérent avec la mission de la DG Marché Intérieur de la Commission Européenne dont dépendait le bureau en charge de Solvabilité 2.

¹³⁴ Voir par exemple leur citation de Mingo (2000) défendant le recours aux modèles internes : « *En effet, si une banque doit supporter un coût en capital supérieur à celui qui découle de son processus interne de mesure des risques à cause des exigences réglementaires, elle ne pourra vraisemblablement pas maximiser la valeur actionnariale* ».

toute conscience, sur la base d'une fidélité théorique, de retenir *dans un objectif de protection des assurés* une grille de lecture utile aux investisseurs et non pertinente pour les assurés.

La deuxième hypothèse est celle d'un choix inconscient, progressif. De fait, François (2015) met certes en évidence l'importance d'une matrice intellectuelle dans la définition des grandes orientations de Solvabilité 2, mais montre également à quel point les options sont restées ouvertes initialement, et comment les choix techniques semblent avoir progressivement échappé à tout contrôle (des expressions telles que « *le monstre* » ou « *les actuaires en folie* » revenant à plusieurs reprises dans les entretiens qu'il cite). Il est ainsi envisageable que les choix techniques décrits *supra* résultent d'une dérive progressive de la complexité conduisant à couper progressivement les questions techniques de leurs enjeux réels (la protection de l'assuré, le contexte d'une entreprise en faillite, etc.) et obligeant, pour trancher entre les différentes options, à se raccrocher uniquement à la cohérence interne, naturellement portée par la logique conceptuelle, donc par une approche financière.

Enfin, au-delà de l'influence des régulateurs dans la définition de la régulation, il pourrait être intéressant de mieux comprendre le poids des entreprises dans ces choix. De fait, l'assimilation de la gestion du risque à la gestion du capital est, au mieux, un levier de transformation de la position du CRO suivant « *une conception du risk management qui est positive [...] au service de la création de richesse* » (Power, 2009) et, au pire, un moyen d'absorber la charge réglementaire et d'en phagocytter les contraintes au sein d'une mission qui sera, et aurait dû être quoi qu'il en soit, gérée en interne.

Savoir que faire

Les enseignements dégagés de cette analyse (alignement des mentalités, dissolution des dirigeants et disparition du risque du cadran d'analyse) appellent, afin d'éviter les effets négatifs résultant des choix retenus et de leurs conséquences, aux orientations suivantes :

- Confier la responsabilité du calcul et du pilotage des métriques dites prudentielles (notamment exigence de capital et ratio de solvabilité) aux CFO et non aux CRO. De fait, ces grandeurs relèvent d'une forme de contrôle de gestion et de la communication financière.
- Pour leur permettre d'exercer un rôle de vigilance face au danger, repositionner les CRO sur l'analyse critique de la stratégie (ouverture du champ des possibles, Frezal, 2015a) plutôt que sur de tels indicateurs qui, *de facto*, sont des indicateurs de rentabilité.
- Veiller notamment à ce que le pilier 2 et le rapport ORSA, que les entreprises ne se sont pas encore approprié, ne soient pas centrés sur une projection du ratio de solvabilité mais correspondent à une analyse des risques extraite de cette grille de pensée imposée.^{135 136}
- Ne pas valoriser, bien au contraire, les *use-tests* qui visent à maximiser la diffusion de ces outils dans les prises de décision (allocation d'actifs, choix de programmes de réassurance, lancement de produits, etc.).

Enfin, il sera souhaitable de tenter de s'extraire de l'arbitrage réglementaire généré par ces pratiques. De fait, l'alignement de la réglementation sur les pratiques de marché, loin de le supprimer, lui a ouvert un nouveau champ en facilitant l'apparition de « coins » entre les nécessaires conventions d'une réglementation et les interprétations alternatives possibles (e.g. sur la frontière des contrats). L'incitation fournie par Solvabilité 2 à la construction de nouveaux produits financiers fondés sur des portefeuilles d'assurance, par exemple les pratiques de titrisation de VIF permettant aux assureurs de réduire leurs SCR, si elle peut représenter l'espoir d'une dispersion des risques au sein des marchés financiers, peut difficilement ne pas faire penser aux produits à

¹³⁵ Interprétation de l'ORSA actuellement portée par certains acteurs, voir par exemple Dacorogna (2015), p12.

¹³⁶ Le danger est double : en substituant aux interdictions de Solvabilité 1 (e.g. interdiction d'investir plus de 40% de ses placements en immobilier) une obligation de penser de telle façon, on impose tout à la fois individuellement un carcan à la pertinence douteuse, et collectivement un appauvrissement réduisant la résilience : chacun pense désormais le risque de la même façon, avec les mêmes outils.

l'origine de la crise de 2008. Un assureur expliquait ainsi à son autorité de contrôle qui lui demandait pourquoi titriser un portefeuille dont le risque était parfaitement maîtrisé :

Oui, à court terme, nous perdons de l'argent en abandonnant une com pour rien. Mais nous habituons le marché, comme les banques ont fait en commençant par titriser le risque de carte de crédit, parfaitement maîtrisé. Et ensuite seulement, elles se sont mises à titriser du risque immobilier. Nous, on fait pareil, on habitue le marché. Et ensuite, on transférera des vrais risques.

Que comprendront les acheteurs de ces produits ? Comment considérer que les risques leur seront effectivement transférés si le vendeur de couverture, à l'instar d'AIG ayant vendu des CDS, n'ayant pas compris ce qu'il acceptait, tombe lui-même en faillite ? Ne pas faire de parallèle semble difficile et il serait audacieux de considérer, dès lors, que les assurés seront correctement protégés par une hypothétique potentialité de transfert vers un reprenneur de l'engagement pris vis-à-vis d'eux.

Dans quels fonds propres pourra-t-on alors puiser pour parvenir malgré tout à indemniser *a minima* les assurés ? Dans les perspectives de profits futurs modélisées. Ici encore, il semble difficile de ne pas faire de parallèle, cette fois-ci avec la méthodologie comptable et les dérives associées ayant largement contribué à la faillite la plus retentissante du début des années 2000...

Références

ACAM (Autorité de Contrôle des Assurances et des Mutuelles), 2006, *Rapport 2006* : https://acpr.banque-france.fr/fileadmin/user_upload/acp/publications/rapports-annuels/ACAM_Rapport_2006_435.pdf

ACPR (Autorité de Contrôle Prudentiel et de Résolution), 2014a, conférence du 18 décembre 2014, https://acpr.banque-france.fr/fileadmin/user_upload/acp/Communication/Conferences/20141218-presentation.pdf

ACPR, 2014, Les chiffres du marché français de l'assurance 2014b, https://acpr.banque-france.fr/fileadmin/user_upload/acp/publications/rapports-annuels/20151007_Rapport_chiffre_2014_Assurances.pdf

Allianz, 2008 à 2014, Allianz Market Consistent Embedded Value Report, https://www.allianz.com/en/investor_relations/results_reports/embedded-value-report.html/

Baud, C., 2013, Le crédit sous Bâle II : un dispositif néolibéral de financiarisation en pratiques, Thèse de doctorat, HEC Paris.

Baud, C. et Chiapello, E., 2015, Disciplining the Neoliberal Bank: Credit Risk Regulation and the Financialization of Loan Management, *Working Paper Series*, <http://papers.ssrn.com/>

Chiapello, E., 2005, Les normes comptables comme institution du capitalisme. Une analyse du passage aux IFRS en Europe à partir de 2005, *Sociologie du travail*, Juillet-Septembre 2005, vol 47, n°3, pp. 362-382.

Dacorogna, M., 2015, A change of paradigm for the insurance industry, *SCOR papers*, n°34.

Deloitte, 2014, Market Consistent Embedded Value, Solid growth amid a changing regulatory landscape, http://www2.deloitte.com/content/dam/Deloitte/es/Documents/servicios-financieros/Deloitte-ES-Servicios-Financieros-market_consistent.pdf

Foulquier, P. (2009), Solvabilité II : une opportunité de pilotage de la performance des sociétés d'assurance, *EDHEC Financial Analysis and Accounting Research Centre* http://professoral.edhec.com/_medias/fichier/edhec-publication-solvabilite-ii-opportunit-e-de-pilotage-swiss-re_1433163287621-pdf

EIOPA (European Insurance and Occupational Pensions Authority), 2006, QIS 1 Summary report, <https://eiopa.europa.eu/Publications/QIS/CEIOPS-FS-0106Rev32006-03-17PA.pdf>

EIOPA, 2008, CEIOPS' Report on its fourth Quantitative Impact Study (QIS4) for Solvency II, <https://eiopa.europa.eu/Publications/QIS/CEIOPS-SEC-82-08%20QIS4%20Report.pdf>

EIOPA, 2013, EIOPA Final Report on Public Consultations No. 13/011 on the Proposal for Guidelines on the Preliminary Application for Internal Models https://eiopa.europa.eu/Publications/Consultations/EIOPA-13-416_Final_Report_on_CP11.pdf

EIOPA, 2014, IORP II in a Nutshell: What is the Holistic Balance Sheet? [https://eiopa.europa.eu/Publications/Reports/7_EIOPA-14-108-IORP II in a Nutshell What is the Holistic Balance Sheet.pdf](https://eiopa.europa.eu/Publications/Reports/7_EIOPA-14-108-IORP_II_in_a_Nutshell_What_is_the_Holistic_Balance_Sheet.pdf)

European Commission, 2010, QIS5 Technical Specifications, http://archive.eiopa.europa.eu/fileadmin/tx_dam/files/consultations/QIS/QIS5/QIS5-technical_specifications_20100706.pdf

European Council, 1973, First Council Directive 73/239/EEC of 24 July 1973 on the coordination of laws, regulations and administrative provisions relating to the taking-up and pursuit of the business of direct insurance other than life assurance, *Official Journal L 228*, 16/08/1973 P. 0003 - 0019

European Council, 1992, Council Directive 92/49/EEC of 18 June 1992 on the coordination of laws, regulations and administrative provisions relating to direct insurance other than life assurance and amending Directives 73/239/EEC and 88/357/EEC (third non-life insurance Directive), *Official Journal L 228*, 11.8.1992, p. 1–23

European Parliament and Council, 2002, Directive 2002/83/EC of the European Parliament and of the Council of 5 November 2002 concerning life assurance, *Official Journal of the European Union* L 345, 19.12.2002, p. 1–51

European Parliament and Council, 2009, Directive 2009/138/EC of the European Parliament and of the Council of 25 November 2009 on the taking-up and pursuit of the business of Insurance and Reinsurance (Solvency II), *Official Journal of the European Union*, L 335-1, 17 December 2009

François, P., 2015, La fabrique d'une tabula rasa : l'émergence de Solvency II, *working paper PARI*, <http://www.chaire-pari.fr/wp-content/uploads/2015/11/Fran%C3%A7ois-Tabula-rasa.pdf>

Frezal, S. 2015a, L'amalgame tyrannique : aléa vs. hétérogénéité, *Working paper PARI*, <http://www.chaire-pari.fr/publications/#working-papers>

Frezal, S., 2015b, Le couple risque / rendement : une chimère ?, *Working paper PARI*, <http://www.chaire-pari.fr/publications/#working-papers>

Frezal, S. 2016, Solvabilité 2 n'est pas risk based, *Working paper PARI*, <http://www.chaire-pari.fr/publications/#working-papers>

Power, M., 2009, The risk management of nothing, *Accounting, Organizations and Society* n°34, pp 849-855

Richard, J., 2015, The dangerous dynamics of modern capitalism (from static to IFRS' futuristic accounting), *Critical Perspectives on Accounting*, Volume 30, July 2015, Pages 9–34

Schleifer, A. et Vishni, R., 1992, Liquidation Values and Debt Capacity: A Market Equilibrium Approach, *The Journal of Finance*, Vol. 47, No. 4 (Sep., 1992), pp. 1343-1366

Conclusion Générale

Apport à la littérature

Cette thèse, par ses apports empiriques, a permis d'étendre au secteur de l'assurance la documentation des dysfonctionnements des outils statistiques pour appréhender les risques, notamment sur quatre points : des marges d'erreur trop importantes pour que de tels outils soient opérationnellement fiables (essai 1) ; des mesures myopes quant aux risques de long terme, et partant susceptibles de fournir des incitations contreproductives (essai 2) ; des effets pervers en termes de gestion, notamment sur la (dé)responsabilisation des acteurs et la prise de décision en (in)conscience (essai 3) et, enfin, la non-pertinence du concept de « juste prix » du risque sur lequel reposent certaines réglementations (essai 4).

Ces quatre premiers essais, au-delà des faits empiriques qu'ils documentent, marquent également des avancées théoriques. Le premier, parce qu'il n'a pas abordé la marge d'erreur sous l'œil extérieur, et exclusivement technique, de la sensibilité du résultat final à un paramètre donné mais, de façon internaliste et holistique quant aux sources potentielles d'erreurs, sous l'angle de la cohérence de différentes estimations produites par un ou des acteurs reconnus du jeu et de la mise en regard de l'information captée et du bruit qui la pollue. Cette approche, qui a ses faiblesses (elle ne permettra jamais de valider la pertinence opérationnelle d'un système de mesures), a aussi ses forces (en l'occurrence, elle permet de l'invalider irréfragablement). Le deuxième, car cela nous a conduit à expliciter et modéliser un mécanisme de procyclicité, spécifique à l'assurance vie, qui n'avait, à notre connaissance, pas encore été formalisé dans la littérature sur la procyclicité des régulations financières, littérature focalisée sur le secteur bancaire et/ou son interaction avec le reste de l'économie. Qui plus est, ce chapitre aura été l'occasion de centrer l'analyse des incitations sur le capital à immobiliser, outil de dialogue entre dirigeants et actionnaires, plutôt

que sur l'exigence de capital, systématiquement considérée dans la littérature bien que somme toute secondaire comme outil de pilotage. Le troisième et le quatrième, car ils ont permis de formaliser le champ de (non-)pertinence des outils statistiques pour appréhender et gérer un phénomène aléatoire, en y introduisant la position de l'observateur ; en identifiant alors les conséquences générales qui découlent de leur utilisation en dehors de ce champ de pertinence, le troisième a permis d'unifier et de justifier les dysfonctionnements documentés à travers la littérature ; Parallèlement, le quatrième a montré que, ainsi qu'une exigence de capital et les coûts d'immobilisation associés – bien que potentiellement pertinents pour un régulateur – ne peuvent refléter, du point de vue d'une compagnie, le « juste coût » dans l'absolu des risques qu'elle prend, de même, une grandeur statistique – bien que potentiellement pertinente pour un assureur – ne peut refléter, du point de vue d'un assuré, le « juste prix » de son risque.

En nous intéressant à l'institutionnalisation de ces outils au sein du secteur de l'assurance, nous avons montré (essai 5) qu'on y retrouve des phases et des vecteurs déjà identifiés dans d'autres domaines. Une phase d'appropriation où certains acteurs au contact d'autres sphères du secteur financier s'acculturent à un cadrage cognitif qui circule alors progressivement vers le secteur de l'assurance. Puis, une phase de mobilisation, où la réglementation joue ici un rôle central, qui crante et diffuse ce cadrage. Enfin, nous montrons que les vertus pratiques de ces mesures, identifiées par la littérature comme parfois unique raison de recours à ces outils, parachèvent ici leur institutionnalisation en conduisant certains des acteurs à les endosser en y adhérant parfois pleinement.

Au-delà de fournir de nouvelles illustrations de motifs (cadrage cognitif, vertus pratiques) et de temporalités (appropriation, mobilisation) déjà identifiés dans la littérature, cet essai présente un apport théorique : appréhender ces outils, leurs dysfonctionnements, leur usage et leur institutionnalisation par le prisme

extrêmement transversal et généralisable de l'amalgame entre les situations dites alea et les situations dites plurielles.

Le dernier essai, en se focalisant sur le design technique de la réglementation, permet d'étendre notre compréhension du pourquoi et comment cette institutionnalisation se produit. Il permet de contraster certaines observations réalisées dans le secteur bancaire, et d'en généraliser d'autres. Ainsi, avec Solvabilité 2, les régulateurs ne sont pas apparus comme un frein à l'émergence d'une nouvelle culture financière, à l'instar de ce qui a pu être documenté sur les marchés financiers, mais au contraire comme un moteur d'une transition culturelle tendant à centrer davantage le pilotage sur la valeur actionnariale. Cette évolution s'est faite comme en banque par un englobement de la finalité de protection des assurés par une logique marchande qui est parvenue à fusionner voire absorber la logique civique. Une telle bascule s'est matérialisée par un renversement des priorités ayant présidé aux choix des caractéristiques techniques de la réglementation, renversement qui avait également été auparavant observé en banque.

Perspectives opérationnelles

Solvabilité 2 a fondé son pilier quantitatif sur une ambition de justesse, par opposition à la vision comptable antérieure, réputée conventionnelle et biaisée. En matière bilancielle, l'objectif n'a pas été atteint. *In fine*, l'ambition combinée de *principle based* et d'exactitude s'est traduite par une dérive de la complexité matérialisée par un nombre incommensurable de conventions. Faute de réalisme dans ce domaine, l'excessive complexité qui s'ensuit est source d'inintelligibilité et de perte de maîtrise du système, tant par ses concepteurs – « experts » – que par ses utilisateurs – « décideurs » –, dirigeants d'entreprise ou régulateurs. Du point de vue des mesures de risque et des exigences de capital associées, l'objectif n'a pas été atteint non plus : les mesures de risque extrême ne sont pas opérationnellement fiables. De ce point de vue, rappelons que le

grief qu'on peut adresser aux calibrages de Solvabilité 2 ne réside pas dans leur caractère imparfait, mais bien dans leur *incapacité* à caractériser les ordres de grandeur relatifs des risques et à les hiérarchiser. Vu l'amplitude des marges d'erreur, l'ambition de justesse à laquelle la simplicité a été sacrifiée est donc chimérique. Chimérique et contreproductive puisque, en substituant à la perception de chacun une norme homogène de perception des risques et de pilotage en incertain, elle tend à aligner les comportements, mettre les décisions en résonance et, de ce fait, accroître le risque systémique, le seul qu'une puissance publique puisse avoir des difficultés à gérer.

Les trois premiers essais pourraient démoraliser tant ils établissent, individuellement et collectivement, à quel point Solvabilité 2, dans sa philosophie même plus encore que dans ses caractéristiques précises, est dans une impasse. Mais, si le constat est effectivement univoque, la situation n'est pas dénuée d'alternative. Paradoxalement, les deux derniers essais tendent à susciter plus d'abattement encore, puisqu'ils montrent que la conscience de dysfonctionnements profonds et inéluctables, directs et indirects, n'est pas nécessairement de nature à remettre en cause l'usage de tels outils. Qui plus est, la littérature nous montre qu'il s'agit là d'un trait généralisé.

Que l'espoir nous soit malgré tout permis : je quitterai désormais le champ de l'analyse pour celui de la recommandation, en suggérant quelques pistes de nature à améliorer le fonctionnement de la régulation prudentielle des assurances. Ces éléments sont proposés dans Frezal (2017), essai en révision dans *Risk Management and Insurance Review*. En raison de son caractère de revue de littérature à visée opérationnelle, centrées sur les conséquences en termes de recommandations pratiques, il n'a pas été associé à cette thèse, mais j'en présente ici les points saillants.

D'un point de vue fondamental, le caractère conventionnel du bilan et de la mesure de risque étant inéluctable, l'enjeu est de choisir la convention opérationnellement la plus saine et non celle théoriquement la plus « juste ».

Par saine, nous entendons la conjonction de trois caractéristiques : (i) simple ; (ii) dont la dimension politique, si elle est présente, est assumée ; (iii) qualitativement pertinente.

- (i) *Simple.* Le caractère simple des outils quantitatifs de valorisation du bilan et de détermination des exigences de capital est essentiel pour le régulateur, pour le marché, et pour chaque entreprise. Pour le régulateur car, disposant de ressources limitées, il s'agit d'une condition nécessaire pour lui permettre d'effectuer un contrôle effectif. Pour le marché car il s'agit d'une condition nécessaire pour limiter les optimisations incontrôlées et donc assurer un *level playing field*. Pour chaque entreprise car il s'agit d'une condition nécessaire pour que les outils de pilotage restent intelligibles aux décideurs et que ceux-ci puissent donc piloter en conscience de leurs limites et de leurs biais.
- (ii) *Dont la dimension politique, si elle est présente, est assumée.* Solvabilité 2 a montré que les enjeux de nature non strictement prudentiels, initialement laissés à distance par ses concepteurs, (financement des marchés action, survivance d'une offre retraite, etc.), sont revenus en boomerang. La double conséquence de la négation officielle de ces enjeux aura donc finalement été :
- au niveau micro-économique, de créer un sentiment fallacieux d'équivalence entre des exigences financières, réglementaires, et des mesures de risque, scientifiques. D'où une perception d'autant plus faussée des risques et une illusion de contrôle,
 - au niveau macro-économique, de ne pas assumer la réflexion sur l'arbitrage entre les enjeux, et ainsi de se couper d'une analyse structurée et d'une décision prise suivant une gouvernance légitime.
- La dimension politique, si elle est présente (et elle le sera nécessairement), doit donc être assumée.

- (iii) *Qualitativement pertinente.* Si les évaluations quantitatives des risques extrêmes ne sont pas fiables, il est en revanche possible, d'un point de vue qualitatif, d'identifier des paramètres accroissant l'incertitude. Par exemple, le fait que le risque est une fonction croissante de la durée des engagements. Il est donc possible d'intégrer, de façon forfaitaire, cette dimension dans les exigences de capital.

Qui plus est, la simplicité des outils et le caractère politique (donc non technique) assumé placeraient les dirigeants en position de développer leur propre analyse : ceci permettrait de limiter l'impact de la réglementation comme vecteur d'homogénéisation des représentations et des décisions qui s'ensuivent, et donc comme source de risque systémique.

D'un point de vue concret, ces éléments fondamentaux se déclinaient prioritairement de la façon suivante :

- (1) Concernant les risques d'actifs, prendre acte de l'incapacité à hiérarchiser les risques rares et ne pas cadrer les comportements d'investissement dans un moule homogène : appliquer un choc de perte de valeur homogène pour les différentes classes d'actif (action, immobilier, crédit) ;
- (2) Concernant les autres risques, basculer sur des calibrages forfaitaires, *uncertainty based*, reflétant qualitativement la hiérarchie des risques (e.g. croissant lorsque la durée des passifs croît), et issus, de façon reconnue et assumée, d'arbitrages politiques avec les enjeux extra-prudentiels ;
- (3) Concernant le bilan, et plus précisément la valorisation des passifs et des fonds propres des compagnies d'assurance vie, substituer aux calculs stochastiques qui président actuellement à la répartition des profits futurs entre actionnaires et assurés un prorata de l'excès d'actifs sur les passifs garantis.

Ces trois éléments, moins techniques qu'ils ne peuvent sembler au premier abord, marqueraient une simplification considérable du système prudentiel et,

par la rupture avec l'illusion de connaissance des risques générée par le système actuel, favoriseraient une appropriation plus personnalisée des risques. Ils sont détaillés et complétés dans Frezal (op.cit.).

Références Générales

Adamko, P, Spuchřáková, E. and Valášková, K., 2015, The History and Ideas Behind VaR, *Procedia Economics and Finance* 24, 18-24

Allais, M., 1953, Le comportement de l'homme rationnel devant le risque : critique des postulats et axiomes de l'école américaine, *Econometrica*, 21 (4), 503-546

Amenc, N., Martellini, L., Foulquier, P. & Sender, S., 2006, The Impact of IFRS and Solvency II on Asset-Liability Management and Asset Management in Insurance Companies, *Edbec Risk and Asset Management Center publication*, 215p.

Armatte, M., 2010, Statut de la dispersion : de l'erreur à la variabilité, *Journal électronique d'histoire des probabilités et de la statistique*, vol. 6 (1)

Baud, C. & Chiapello, E., 2017, Understanding the disciplinary aspects of neoliberal regulations: The case of credit-risk regulation under the Basel Accords, *Critical Perspectives on Accounting*, Volume 46, July 2017, 3-23

Black, F. & Scholes, M., 1973, The Pricing of Options and Corporate Liabilities, *Journal of Political Economy* 81.3, 637-654

Boltanski, L. & Thévenot, L., 2006, *On justification. Economies of worth*. Princeton university press, Princeton

Callon, M. & Muniesa, F., 2005, Economic markets as calculative collective devices, *Organization Studies*, 26 (8), 1229-1250

- Camerer, 1989, An experimental Test of Several Generalized Utility Theories, *Journal of Risk and Uncertainty*, 2, 61-104
- Chiappello, E. & Gilbert, P., 2009, La gestion comme technologie économique, in Steiner, P., Vatin, F. (dir.), *Traité de sociologie économique*, PUF, Paris
- Chiappello E., 2015, Financialisation of Valuation, *Human Studies*, 38 (1), 13-35
- Consumer Federation of America (CFA), 2013, Letter sent 29 August 2013 by R. Hunter, Director of Insurance, Casualty Actuarial Society.
- Court of Justice of the European Union (CJEU), 2011, 1 March 2011 Judgment in case C-236/09, curia.europa.eu.
- Danielsson J., 2008, Blame the models, *Journal of Financial stability*, 4 (4). 321-328
- Danielsson J., 2002, The emperor has no clothes: Limits to risk modelling, *Journal of Banking and Finance*, vol. 26, issue 7, 1273-1296
- Daston, L., 1988, *Classical Probability in the Enlightenment*, Princeton University Press, New Jersey.
- Daston, L. 1989, L'interprétation classique du calcul des probabilités, *Annales, Economies, Sociétés, Civilisations*, An. 44, N°3, 715-731.
- Doherty, N., 2013, Risk and the Endogenous Economist: Some Comparisons of the Treatment of Risk in Physics and Economics, *Geneva Risk and Insurance Review*, 38 (1), 1-22.
- Eagle, A. 2004, Twenty-One Arguments Against Propensity Analyses of Probability, *Erkenntnis*, 60, 371-416.
- El Karoui, N., Loisel, S., Prigent, J.-L., & Vedani, J., 2015, Market inconsistencies of the market-consistent European life insurance economic valuations: pitfalls and practical solutions, *working paper*, <https://hal.archives-ouvertes.fr/hal-01242023>

Ellsberg, D., 1961, Risk, Ambiguity, and the Savage Axioms, *Quarterly Journal of Economics*. 75 (4), 643–669.

Espeland, W. and Stevens, M., 1998, Commensuration as a Social Process, *Annual Review of Sociology*, 24:1, 313-343

European Commission (EC), 2009, Directive 2009/138/EC of the European Parliament and of the Council on 25 November 2009 on the taking-up and pursuit of the business of Insurance and Reinsurance (Solvency II), *Official Journal of the European Union*, L335/1, 17.12.09.

European Commission (EC), 2011, *Press Release* - European Commission gives guidance to Europe's insurance industry to ensure non-discrimination between women and men in insurance premiums. Available online at: http://europa.eu/rapid/press-release_IP-11-1581_en.htm

European Council (EC), 1992, Council Directive 92/49/EEC of 18 June 1992 on the coordination of laws, regulations and administrative provisions relating to direct insurance other than life assurance and amending Directives 73/239/EEC and 88/357/EEC (third non-life insurance Directive), *Official Journal L 228*, 11.8.1992, p. 1–23

European Council, (EC), 2004, Directive 2004/113/EC Implementing the Principle of Equal Treatment between Men and Women in the Access to and Supply of Goods and Services. *Official Journal of the European Union*, L373 1, 21.12.04.

European Parliament and Council (EPC), 2002, Directive 2002/83/EC of the European Parliament and of the Council of 5 November 2002 concerning life assurance, *Official Journal of the European Union L 345*, 19.12.2002, p. 1–51

Fédération Française de l'Assurance (FFA), 2017, Tableau de bord de l'assurance 2016

- Feller, W., 1968, *An Introduction to Probability Theory and its Applications*, Vol. 1, 3rd ed., John Wiley & Sons Inc., New York.
- Foulquier, P., & Arias, L., 2012, Solvabilité 2 : l'approche par les risques, *Revue Banque*, n°745
- François, P., 2011, Puissance et genèse des institutions : un cadre analytique, in François, P. (dir.), *Vie et mort des institutions marchandes*, Paris, Presses de sciences po, 39-79.
- François, P., 2015, La fabrique sociale d'une tabula rasa. Le lancement de Solvency II, *Working paper PARI*
- Frezal, S., 2018, Solvency II Is Not Risk-Based—Could It be ? Evidence from Non-Life Calibrations », *North American Actuarial Journal*, 22(1), 15p.
- Frezal, S., Haguët, E., and Nou, V., 2016, With-Profit Contracts Regulation and Procyclicality: A Comparative Analysis of Alternative Approaches, *Working paper PARI*, 31p.
- Frezal, S., 2016, Risk and statistics: a regulation fallacy, *Working paper PARI*, 35p.
- François, P., et Frezal, S., 2018, Instituer l'incohérence : aléa et hétérogénéité au sein du secteur assurantiel, *Sociologie du travail*, Vol. 60 – n° 1.
- Frezal, S., 2015, De quoi Solvabilité 2 est-il le nom ?, *Working paper PARI*, 56p.
- Frunza, M., 2013, Impact de Solvabilité II sur l'économie réelle : une approche micro-économique, *Working paper Labex ReFi*
- Gilboa, I., Postlewaite, A. and Schmeidler, D., 2008, Probability and Uncertainty in Economic Modeling, *Journal of Economic Perspectives*, Vol 22, N°3, 173-188.
- Gilboa, I. 2009, Is it always rational to satisfy Savage's axioms?, *Economics and philosophy*, 25(3), 2285-296

- Gollier, C., Hammitt, J. K., & Treich, N., 2013, Risk and choice: A research saga, *Journal of risk and uncertainty*, 47(2), 129-145
- Guibert, Q., Julliard, M., Nteukan-Teugua, O., & Planchet, F., 2014, *Solvabilité prospective en assurance*, Economica, Paris
- Haldane A. and Madouros V., 2012, The Dog and the Frisbee (speech), *Bank of England*
- Harrison, M. & Samuel, S., 2011, Two hundreds millions inputs. Can you trust risk weightings at European banks?, *Barclays Capital Equity Research*, 06.04.2011
- Huault, I. et Rainelli, H., 2007, Innovations financières : Construire et légitimer un nouveau marché financier de gré à gré - le cas des dérivés de crédit, *Finance Contrôle Stratégie*, Vol. 10, n° 1
- Huault, I. et Rainelli, H., 2009, Market Shaping as an Answer to Ambiguities: the case of credit derivatives, *Organization Studies*, Vol. 30, n° 5, 549-575
- Huault, I. et Rainelli, H., 2011, A Market for weather risk ? Conflicting metrics, attempts to compromise and limits to commensuration *Organization Studies*, Vol. 32, n° 10
- Huault, I. et Rainelli, H., 2016, Business as usual in Financial Markets? The creation of incommensurables as institutional maintenance work, *Organization Studies*, Vol. 37, n°7, 991-1015
- Hubert, H., Mauss, M., 1985, Esquisse d'une théorie générale de la magie, in Mauss, M., *Sociologie et anthropologie*, Presses universitaires de France, Paris, 3-142
- Jones, D., 2013, Re: Statement of principles Regarding Property and Casualty Insurance Ratemaking. *Letter to the Chief Actuary of Louisiana Department of Insurance*, 2013, May 21.
- Kahneman, D. and Tversky, A., 1979, Prospect Theory: An Analysis of Decision under Risk, *Econometrica*, Vol. 47, No. 2, 263-292

- Kamega, A. & Planchet F., 2013, Construction de tables de mortalité prospectives sur un groupe restreint : mesure du risque d'estimation, Bulletin français d'Actuariat, *Bulletin Français d'Actuariat*, vol. 13, n°25
- Kashyap, A. and Stein, J., 2004, Cyclical implications of the Basel II capital standards, *Economic Perspectives*, Q1, 18-31
- Knight F., 1921. *Risk, Uncertainty and Profit*, Houghton Mifflin, Boston and New York
- McKenzie, D. et Millo, Y., 2003, Constructing a market: the historical sociology of a financial derivatives exchange, *American journal of sociology*, vol. 109 (1), 107-145
- MacKenzie, D. and Millo Y., 2009, The Usefulness of Inaccurate Models: The Emergence of Financial Risk Management, *Accounting, Organizations and Society* no. 34, 638-653
- Mandelbrot, B. & Hudson, L., 2009, *Une approche fractale des marchés. Risquer, perdre et gagner*, Paris, Odile Jacob
- Markowitz, H., 1952, Portfolio selection, *The journal of finance*, vol. 7, n°1, 77-91
- McGoun E., 1995, The history of risk “measurement”, *Critical Perspectives on Accounting*, 6, 511-532
- Mikes, A., 2009, “Risk management and calculative cultures”, *Management Accounting Research*, 20(1), 18-40
- Mikes, A. 2011, “From Counting Risk to Making Risk Count: Boundary-Work in Risk Management”, *Accounting, Organizations and Society*, 36, 226-245
- Miller, M. J., 2009, “Disparate Impact and Unfairly Discriminatory Insurance Rates.” *Casualty Actuarial Society E-Forum*, n° Winter 2009.
<https://www.casact.org/pubs/forum/09wforum/>.

- Münzer, M. 2016, Investigating tensions between competing institutional logics and orders of worth in regulatory post-crisis decision-making – the case of the French structural banking reform, in *Financial market regulation in the aftermath of the financial crisis – three essays on structural banking reforms*, PhD thesis, Université de Strasbourg
- Observatoire de l'évolution des métiers de l'assurance (OEMA), 2017, Synthèse ROMA/ROFA 2017
- Pfeifer, D. and Strassburger, D., 2008, "Solvency II: Stability problems with the SCR aggregation formula", *Scandinavian Actuarial Journal*, 1, 61-77
- Planchet, F. & Théron, P., 2007, *Mesure et gestion des risques d'assurance*, Economica, Paris
- Plantin, G., Sapra, H. and Shin, H., 2008, Marking to market: Panacea or Pandora's Box", *Journal of Accounting Research*, 46, 435-60
- Porter, T. M., 1986, *The rise of statistical thinking, 1820-1900*, Princeton university press.
- Power, M., 2009, The risk management of nothing, *Accounting, organizations and society*, vol. 34, 849-855.
- Roe M., 1994, *Strong Managers, Weak Owners. The Political Roots of American Corporate Finance*, Princeton University Press
- Samuelson, P., 1963. Risk and Uncertainty: A Fallacy of Large Numbers, *Scientia*, 98, 108-113
- Taylor, A. and Goodhart, C., 2004, Procyclicality and volatility in the financial system: the implementation of Basel II and IAS 39, *London School of Economics, FMG Discussion Paper*.
- Thévenin, L. & Vial, G., 2009, Solvabilité II : plus que deux ans pour se mettre en conformité, *Les Echos*, 21/10/2009

- Von Mises, L., 1949. *Human Action*, I, VI, 5, New Haven, Yale University Press
- Von Neumann, J. and Morgenstern, O., 1944. *Theory of Games and Economic Behaviour*, Princeton University Press
- Walter, C. & Pracontal, M., 2009, *Le Virus B. Crises financières et mathématiques*, Paris, Seuil.
- Walter, C., 2013, Les origines du modèle de marche au hasard en finance, *Working paper*, halshs-00828289
- Savage, L., 1954. *The Foundations of Statistics*, New York, Wiley
- Sims, C., 2003, Implications of rational inattention, *Journal of Monetary Economics*, Volume 50, Issue 3, April 2003, 665–690
- Sharpe, W. F., 1964, Capital asset prices: A theory of market equilibrium under conditions of risk, *Journal of Finance*, 19 (3): 425-442
- Taleb N. N. 2007, *The Black Swan*, Random House
- Zelizer, V., 1979, *Morals and markets*, Columbia university press, New York.

Sommaire Détaillé

Remerciements	3
Introduction Générale	5
1. Positionnement du sujet	5
1.1 Contexte : l'assurance, un secteur crucial profondément impacté par la réforme prudentielle Solvabilité II	5
1.2 Etat de l'art : les statistiques, un outil de gestion des risques qui dysfonctionne	8
1.3 Question de recherche générale	13
2. Questions de recherche	14
2.1 Les limites de l'approche statistique adoptée par Solvabilité 2	14
2.1.1 Le problème de la fiabilité des estimateurs	14
2.1.2 La question de la pertinence des indicateurs	15
2.2 Dysfonctionnements des mesures de risque dans la régulation des assurances : proposition d'un cadre théorique et applications au coût du risque	16
2.2.1 Les défauts de l'approche statistique des risques : proposition de théorisation et application à Solvabilité 2	16
2.2.2 Coût et tarification des risques : de Solvabilité 2 à la question de l'équité des prix des assurances	17
2.3 Retour sur les facteurs d'institutionnalisation de l'approche statistique promue par Solvabilité 2	18
2.3.1 La réussite d'un amalgame problématique	18
2.3.2 Le basculement vers une logique marchande	19
3 Résumé	20
3.1 Partie I : Les limites de l'approche statistique adoptée par Solvabilité 2	21
3.1.1 Chapitre 1 : Le problème de la fiabilité des estimateurs	21
3.1.2 Chapitre 2 : La question de la pertinence des estimateurs	23
3.2 Partie II : Dysfonctionnements des mesures de risque dans la régulation des assurances : proposition d'un cadre théorique et applications au coût du risque	25
3.2.1 Chapitre 3 : Les défauts de l'approche statistique des risques : proposition de théorisation et application à Solvabilité 2	25
3.2.2 Chapitre 4 : Coût et tarification des risques : de Solvabilité 2 à la question de l'équité des prix des assurances	27
3.3 Partie III : Retour sur les facteurs d'institutionnalisation de l'approche statistique promue par Solvabilité 2	28
3.3.1 Chapitre 5 : La réussite d'un amalgame problématique	28
3.3.2 Chapitre 6 : Le basculement vers une logique marchande	30
Partie I : Les limites de l'approche statistique adoptée par Solvabilité 2	32
Chapitre 1 : Le problème de la fiabilité des estimateurs	32
Introduction	33

1. Are these risk measures “better then nothing”?	40
1.1 Methodology	40
Idealistic world	41
Real World	41
1.2 Data	45
1.2.1 QIS	45
1.2.2 An internal model	46
1.3 Results	46
1.3.1 QIS	46
1.3.2 Internal model	47
1.4 Interpretation: is the “noise” truly noise?	48
2. The causes	51
2.1 Political economy?	51
2.1.1 Data used	51
2.1.2 Results	53
2.2 An idiosyncrasy?	54
2.3 An epistemological barrier?	56
Conclusion	58
Better identification of technical limitations	59
Conscientious steering	60
References	61
Appendix	64
<i>Chapitre 2 : La question de la pertinence des indicateurs</i>	70
Introduction	71
1. With-profit contracts, intuition and modeling	74
1.1 Contract description	74
1.2 Modeling of the contract and associated profit function	75
1.3 Valuation of economic capital	76
1.4 Risk-based capital requirement and procyclicality	78
2. Comparative steering insights of risk-based, economic (RBE) system vs. non-RBE system	81
2.1 Minimum raised capital and anticipation of slow risks	82
2.2 Sensitivity of the minimum raised capital and anticipation of short-term risks	85
3. Implementation within Solvency II	87
3.1 The characteristics of the simulation	87
3.2 Capital requirements	89
3.3 Minimum raised capital	90
Conclusion	91
References	93

Partie II : Dysfonctionnements des mesures de risque dans la régulation des assurances : proposition d'un cadre théorique et applications au coût du risque _____ **97**

Chapitre 3 : Les défauts de l'approche statistique des risques : proposition de théorisation et application à Solvabilité 2 _____ **97**

Introduction _____ **98**

1. Underlying conceptual framework and derived general results _____ **100**

1.1 Definitions _____ 100

1.2 Graphic intuition _____ 103

1.3 General results _____ 105

1.4 Summary of the theoretical framework _____ 108

2. Capital requirements _____ **110**

2.1 The regulatory debate _____ 110

2.2 The critical analysis _____ 110

2.3 The operational solution _____ 114

Conclusion _____ **118**

References _____ **121**

Appendix _____ **123**

Appendix A: Legal heuristics which tackle this conflation _____ 123

Appendix B: Mathematical formalization of an *à la Solé* decision making process under uncertainty _____ 124

Chapitre 4 : Coût et tarification des risques : de Solvabilité 2 à la question de l'équité des prix des assurances _____ **131**

Introduction _____ **132**

1. Insurance Fairness and Knowledge _____ **134**

1.1 A Historical Overview of Insurance Fairness _____ 135

1.2 Epistemological Assumptions behind Actuarial Fairness _____ 137

1.3 Two Conceptions of Fairness _____ 141

2. Plurality, Randomness and Actuarial Fairness _____ **143**

2.1 Plurality vs. Randomness: a Conceptual Framework _____ 143

2.2 On expected value as the fairness reference framework _____ 146

2.3 On expected value as an objective measure of the individual "true risk" _____ 148

3. Discussion: Fairness in Insurance _____ **149**

Conclusion _____ **150**

References _____ **151**

Partie III : Retour sur les facteurs d'institutionnalisation de l'approche statistique promue par Solvabilité 2 _____ **159**

Chapitre 5 : La réussite d'un amalgame problématique	159
Introduction	160
1. L'appropriation de l'amalgame par l'industrie financière	168
1.1. L'appropriation des indicateurs de risque, ou l'amalgame au cœur des pratiques	169
1.2. Le calcul stochastique : des produits dérivés à la gestion prudentielle	172
2. Appropriation et mobilisation de l'amalgame dans le secteur assurantiel	175
2.1. Transfert et appropriation	176
2.2. Évaluer les entreprises d'assurance : la mobilisation des CFO	178
2.3. Définir de nouvelles règles du jeu	181
3. L'appropriation reconduite de l'amalgame	187
3.1. Le registre de la foi	187
3.2. Réflexivité et cynisme dans l'usage des institutions	191
4. Conclusion	194
Références	196
Chapitre 6 : Le basculement vers une logique marchande	202
Introduction	203
Solvabilité 2 : quel objectif ?	203
La trame et l'intuition	206
<i>Caveat</i>	207
1. Les faits : description des systèmes prudentiels	208
1.1 La pierre angulaire : le bilan	208
1.2 Une mise en regard usuelle devant être approfondie	210
1.3 Solvabilité 1	212
Provisionnement non vie	212
Provisionnement vie	213
Actifs	214
Fonds propres disponibles	215
Exigences de capital	216
Synthèse	217
1.4 Solvabilité 2	218
Actifs	219
Provisions	219
Fonds propres disponibles	221
Exigences de capital	222
Synthèse	225
2. Pour qui les concepts de Solvabilité 2 sont-ils significatifs ?	227
2.1 En vie, un <i>best estimate</i> ayant du sens pour l'actionnaire, non pour l'assuré	228
L'objectif informationnel n'est plus la protection de l'assuré contre la faillite	230

L'estimation est porteuse de sens pour l'actionnaire et non pour l'assuré _____	231
L'arbitrage biais/marge d'erreur correspond à une logique d'investissement et non de régulation _____	233
Synthèse : calculer le best estimate, une évolution « inutile et incertaine » d'un point de vue prudentiel _____	238
2.2 Un provisionnement global conduisant à définir la faillite d'une façon qui a du sens pour l'actionnaire, non pour l'assuré _____	239
2.3 Des fonds propres et exigences de capital associés à une notion de risque définie d'une façon qui a du sens pour l'actionnaire, non pour l'assuré _____	243
2.4 Quand une crise survient... _____	247
3. Quelques conséquences possibles _____	250
Solvabilité 2, catalyseur pour l'assurance d'un processus économique de fond ____	250
Une logique de disparition des dirigeants ? _____	252
Vers une disparition de la prise en compte du risque ? _____	255
L'état d'esprit des régulateurs est également modifié _____	257
Conclusion _____	258
Synthèse _____	258
Comprendre pourquoi _____	261
Savoir que faire _____	264
Références _____	266
<i>Conclusion Générale</i> _____	270
Apport à la littérature _____	270
Perspectives opérationnelles _____	272
<i>Références Générales</i> _____	276
<i>Sommaire Détaillé</i> _____	284

DE L'USAGE DES STATISTIQUES POUR LE PILOTAGE ET LA REGULATION DES RISQUES EN ASSURANCE

Résumé

L'usage des grandeurs statistiques pour éclairer la décision en situation de risque, apparu au 18ème siècle puis disqualifié au 19ème siècle, a été réintroduit au milieu du 20ème siècle et s'est depuis progressivement imposé dans l'industrie financière, percolant dans l'assurance jusqu'à s'y généraliser *via* Solvabilité 2.

Pourtant, de nombreux dysfonctionnements liés à la mobilisation de ces outils pour la gestion et la régulation des risques ont été mis en lumière par la littérature économique, actuarielle, et en sciences de gestion. Par conséquent, cette extension du champ d'application de tels outils à la régulation des assurances interpelle.

Cette thèse (i) étend empiriquement au secteur de l'assurance la littérature produite dans d'autres secteurs afin d'y identifier les limites de l'utilisation des statistiques pour la gestion des risques et la régulation prudentielle, (ii) propose un cadre théorique unifié pour les dysfonctionnements liés à ces usages, et (iii) éclaire les ressorts de l'adoption de ces outils au sein du secteur de l'assurance.

Risque, régulation, assurance, pilotage, prudentiel, Solvabilité 2

Résumé en anglais

The use of statistics for risk management and decision making under uncertainty, discredited in the nineteenth century, was reintroduced in the mid twentieth century and, since then, widespread progressively in the financial industry up to being generalised into the insurance sector via Solvency II.

However, economic, actuarial and management science literature highlighted numerous dysfunctions of these tools when used for risk management and regulation. As a consequence, such expansion of the field of these tools is compelling.

This thesis (i) empirically expands scientific literature towards the insurance field in order to identify the limits of the use of statistics for risk management and prudential regulation, (ii) proposes a unified theoretical framework of the subsequent dysfunctions, and (iii) sheds lights on the reasons why these tools spread among the insurance sector.

Risk, regulation, insurance, steering, prudential, Solvency II