

**Mémoire présenté devant l'Institut du Risk Management
pour la validation du cursus à la Formation d'Actuaire
de l'Institut du Risk Management
et l'admission à l'Institut des actuaires
le**

Par : Laetitia BADOURALY-KASSIM et Adrian IUGA

Titre : Etude transversale sur le risque de réserve en non-vie
sur le marché français de l'assurance

Confidentialité : NON OUI (Durée : 1an 2 ans)
Les signataires s'engagent à respecter la confidentialité indiquée ci-dessus

Membres présents du jury de l'Institut des
actuaires :

Entreprise : ACPR

Nom : Eric MOLINA

Signature et Cachet :

Eric Molina

Membres présents du jury de l'Institut du Risk
Management :

Directeur de mémoire en entreprise :

Nom : Eric MOLINA

Signature : *Eric Molina*

Invité :

Nom :

Signature :

**Autorisation de publication et de mise en
ligne sur un site de diffusion de documents
actuariels**

(après expiration de l'éventuel délai de confidentialité)

Signature du responsable entreprise

Secrétariat :

Bibliothèque :

Signature(s) du candidat(s)

Iugara *LBadouraly*

MÉMOIRE D'ACTUARIAT
dans le cadre du
CENTRE D'ETUDE ACTUARIELLE
réalisé au sein de
AUTORITÉ DE CONTRÔLE
PRUDENTIEL ET DE RÉOLUTION



Étude transversale du risque de réserve
en non-vie sur le marché français de
l'assurance



Laetitia BADOURALY KASSIM
et
Adrian IUGA

Mémoire encadré par Eric MOLINA

Table des matières

Table des matières	iii
1 Motivation et plan du mémoire	1
2 Notations, méthodes de provisionnement et d'estimation du risque de provisionnement	5
2.1 Provisionnement prudentiel et comptable	6
2.2 Triangle de liquidation	7
2.3 Modèle <i>Chain-Ladder</i>	8
2.4 Modèle de Mack	11
2.5 Modèle de Merz & Wüthrich	12
2.6 Méthode Bootstrap	14
2.7 Lemme pour la réassurance proportionnelle	16
2.8 Lien avec la charge en capital réglementaire Solvabilité 2 en formule standard	18
2.9 Méthode réglementaire pour les paramètres spécifiques	22
2.10 Calibrage de la formule standard par EIOPA	24
3 Méthodologie de travail et comparaison des méthodes	29
3.1 Description des données utilisées	30
3.2 Contrôle de la qualité des données	33
3.3 Algorithme de traitement des données	36
3.4 Retraitement des données	37
3.5 Triangles de charges inexploitable	47
3.6 Différence entre les écarts-types des méthodes de Merz & Wüthrich et Bootstrap	48
3.7 Convergence dans la méthode Bootstrap	49
3.8 Méthode Bootstrap avec d'autres lois de simulation	50
4 Résultats numériques	53
4.1 Segment "autre assurance des véhicules à moteur"	54
4.2 Segment "responsabilité civile automobile"	57
4.3 Segment "responsabilité civile générale"	61
4.4 Segment "incendie et autres dommages aux biens"	64
4.5 Segment "pertes pécuniaires diverses"	67

4.6	Segment “protection juridique”	70
4.7	Segment “frais médicaux”	73
4.8	Segment “protection du revenu”	75
4.9	Segment “assistance”	79
5	Conclusions	81
	Bibliographie	89

CHAPITRE

1

Motivation et plan du mémoire

L'assurance non-vie repose avant tout sur le principe de mutualisation des risques. Les assureurs cherchent ainsi à regrouper des assurés potentiellement exposés à des risques identiques afin de diminuer le risque de ruine et les fonds propres à immobiliser. En pratique, le risque (et donc les fonds propres à immobiliser) ne croît pas de manière linéaire avec l'exposition.

Le modèle théorique le plus simple consiste à considérer que les risques sont indépendants et identiquement distribués. Sous cette hypothèse, l'espérance de la charge à payer par l'assureur est proportionnelle au nombre de sinistres n alors que le risque associé, en termes d'écart-type, croît moins vite et est proportionnel à \sqrt{n} . Le coefficient de variation (défini comme l'écart-type divisé par l'espérance) décroît avec la taille de la population assurée en $\frac{1}{\sqrt{n}}$.

Dans le cadre du régime prudentiel Solvabilité 2, en formule standard pour un segment d'activité donné, l'exigence en capital est proportionnelle à l'exposition, le bien connu "trois sigma V" (i.e. $3 \times \sigma \times V$, avec σ le coefficient de variation donné par le Règlement Délégué et V l'exposition). Ainsi, le régime prudentiel Solvabilité 2 ne capte pas l'effet non-linéaire entre l'exposition et le risque. Cela a amené les plus importants acteurs du marché français de l'assurance non-vie à développer des modèles internes ou à utiliser des paramètres spécifiques calibrés sur les données propres (sous l'accord de l'Autorité de contrôle prudentiel et de résolution).

L'objet principal de ce mémoire est d'analyser des données du marché français de l'assurance afin d'étudier la nature de la décroissance du coefficient de variation (l'équivalent du « sigma » de la formule standard) pour le risque de réserve sur différents segments d'activité Solvabilité 2. Le modèle théorique simplifié où les risques sont indépendants et identiquement distribués revient à une décroissance en loi puissance avec un exposant de $-0,5$. La formule $3 \times \sigma \times V$ revient à considérer une fonction constante (puissance avec un exposant de 0).

L'étude sera menée sur le risque de réserve. D'un point de vue pratique, cela revient à tracer le coefficient de variation ($\frac{\sigma(BE)}{BE}$) en fonction du volume afin d'étudier la vitesse de décroissance.

L'écart-type lié au risque de réserve pourra être calculé en considérant plusieurs méthodes. Par exemple, nous pourrions considérer la méthode de Mack, la méthode de Merz Wuthrich, ou encore envisager le calcul par la technique de *bootstrap*. Les données nécessaires au calcul du coefficient de variation seront directement extraites des états prudentiels Solvabilité 2, et plus précisément des états s.19.01.01. Ces états contiennent, pour chaque société et chaque segment d'activité exercé, l'ensemble des triangles de liquidation.

Dans le cadre de ce mémoire nous allons tester la nature de la décroissance du coefficient de variation, et étudier la forme de la décroissance obtenue en fonction de la mesure de risque choisie et de la ligne d'activité considérée. Si une forme de décroissance stable est obtenue, il pourra ainsi en être déduit le volume de réserve correspondant au coefficient de la formule standard. De manière plus générale, l'adéquation de la formule standard au marché français de l'assurance non-vie sera analysée.

De plus, cette étude permettra de comparer d'une manière transversale les différents segments d'activité sur le marché français, grâce à plusieurs indicateurs statistiques : le temps et la cadence de liquidation des sinistres, la différence entre les modèles théoriques et les approximations mises en place dans la réglementation (par souci de simplification) et les écarts numériques des différents modèles actuariels utilisés pour la modélisation du risque de réserve en non-vie.

Ce mémoire est organisé en trois parties. La première rappelle les modèles actuariels les plus connus pour l'analyse du provisionnement et le calcul du risque de provisionnement. Une attention particulière est portée sur le calibrage de la formule standard, avec une analyse des modèles sous-jacents utilisés pour le calibrage et une analyse de la méthodologie suivie pour l'application de ces modèles. Dans une seconde partie, nous décrivons l'ensemble des données utilisées dans le cadre de ce mémoire et expliquons comment ces données ont été traitées, notamment face aux problèmes de qualité des données. Des écarts entre les différents choix de modélisation sont indiqués afin de sensibiliser le lecteur aux limites de cette étude. Enfin, dans une dernière partie, les résultats numériques sont indiqués et analysés pour chaque segment d'activité, ce qui permettra de conclure sur l'ensemble des problématiques soulevées.

2

Notations, méthodes de provisionnement et d'estimation du risque de provisionnement

2.1	Provisionnement prudentiel et comptable	6
2.2	Triangle de liquidation	7
2.3	Modèle <i>Chain-Ladder</i>	8
2.4	Modèle de Mack	11
2.5	Modèle de Merz & Wüthrich	12
2.6	Méthode Bootstrap	14
2.7	Lemme pour la réassurance proportionnelle	16
2.8	Lien avec la charge en capital réglementaire Solvabilité 2 en formule standard	18
2.9	Méthode réglementaire pour les paramètres spécifiques	22
2.10	Calibrage de la formule standard par EIOPA	24

Cette section est dédiée au rappel de certaines méthodes de provisionnement classiques en assurance non-vie, à l'estimation de l'erreur de provisionnement et au traitement prudentiel du risque de provisionnement.

2.1 PROVISIONNEMENT PRUDENTIEL ET COMPTABLE

Du fait de l'inversion du cycle de production en assurance, les sociétés d'assurance et de réassurance estiment des provisions pour faire face à leurs engagements qui sont inscrites au passif de leur bilan (comptable ou prudentiel). Ces provisions sont le poste le plus significatif du bilan d'un assureur. Dans le cas des entités non-vie et santé non-similaire à la vie, de manière simplifiée, les provisions sont séparées entre les sinistres déjà survenus et les sinistres futurs. La provision la plus importante pour un assureur non-vie est celle concernant **les sinistres déjà survenus et qui font l'objet d'étude de ce mémoire**. Dans le cadre comptable on parle de **la provision pour sinistres à payer ou PSAP** alors que dans le cadre prudentiel on parle de **la meilleure estimation de la provision pour sinistres à payer, nommée aussi "Best Estimate de réserves"** (ou encore *BE de réserves*).

Dans le cadre prudentiel d'aujourd'hui, connu sous le nom de Solvabilité 2 (voir la directive [6]), les provisions en assurance non-vie sont décrites à l'article 36 du règlement délégué [7]. En ce qui concerne la meilleure estimation pour sinistres à payer, elle est définie au paragraphe 3 de l'article cité *supra* comme se rapportant "à des sinistres qui se sont déjà produits, indépendamment de la question de savoir s'ils ont été déclarés ou non". De plus, il est précisé que "les projections de flux de trésorerie pour le calcul de la provision pour sinistres à payer comprennent les prestations, dépenses et primes relatives aux sinistres visés au paragraphe 3."

Ainsi, cette définition est proche de la provision pour sinistres à payer (PSAP) du cadre comptable définie dans le Code des Assurances [3] à l'article R.343-7, paragraphe 4 : "Provision pour sinistres à payer : valeur estimative des dépenses en principal et en frais, tant internes qu'externes, nécessaires au règlement de tous les sinistres survenus et non payés, y compris les capitaux constitutifs des rentes non encore mises à la charge de l'entreprise."

Cependant, trois principales différences existent entre ces deux définitions :

- **actualisation** : Pour la provision Solvabilité 2, en vision meilleure estimation (ou *best estimate*), il est tenu compte de la valeur temporelle de l'argent. Les flux futurs sont ainsi actualisés avec une courbe de taux¹ dite "EIOPA"², ce qui n'est pas le cas de la PSAP comptable ;
- **primes** : Dans le cas des primes relatives à des sinistres déjà survenus et qui sont versées après la date de clôture, celles-ci devrait être comptées dans la meilleure estimation de la provision pour sinistres (avec un signe négatif). C'est notamment le cas des primes acquises non émises (PANE). D'ailleurs, l'Autorité de contrôle prudentiel et de résolution (ACPR) a in-

1. EIOPA publie chaque mois une courbe de taux dont l'utilisation est obligatoire pour l'actualisation de la meilleure estimation des engagements.

2. European Insurance and Occupational Pensions Authority.

diqué que la déduction devrait être affichée dans la meilleure estimation pour primes afin de ne pas diminuer indûment le risque de réserve. Sur le marché français, ce cas n'est pas courant et concerne notamment la prévoyance collective (et à titre particulier les institutions de prévoyance). En comptabilité (référentiel français ou norme IFRS4), ces primes sont inscrites à l'actif comme des créances ;

- **principe de prudence** : Alors que la provision comptable est calculée avec une approche de prudence, la meilleure estimation est calculée dans une approche "50-50".

La méthode réglementaire d'évaluation de la PSAP comptable est celle du dossier-dossier. Pour chaque sinistre déclaré (et pas encore clos), un montant est estimé par un gestionnaire sinistre de ce qui reste à payer. A cela, s'ajoute une estimation pour les sinistres survenus avant la date d'inventaire mais non connus à la date des calculs. On parle dans ce cas des "tardifs" ou IBNR (pour *Incurring But Not Reported*). Depuis 1991, l'autorité de contrôle autorise l'utilisation de méthodes statistiques en utilisant des données historiques. En ce qui concerne la "nouvelle" réglementation Solvabilité 2, d'après l'article 34 du règlement délégué [7], aucune méthode n'est imposée tant que "la meilleure estimation est calculée d'une manière transparente, de telle sorte que la méthode de calcul et les résultats qu'elle produit peuvent être contrôlés par un expert qualifié". De plus, il est précisé que "le choix des méthodes actuarielles et statistiques pour le calcul de la meilleure estimation est fondé sur la capacité de ces méthodes à refléter les risques affectant les flux de trésorerie sous-jacents, ainsi que sur la nature des engagements d'assurance et de réassurance". Cela laisse donc comprendre que les modèles actuariels seraient privilégiés tant que ces "méthodes actuarielles et statistiques sont cohérentes avec toutes les données disponibles".

Dans le cadre de ce mémoire, l'attention sera apportée sur le provisionnement prudentiel dit *best estimate* à partir d'une méthode statistique connue sous le nom de *Chain-Ladder*, largement utilisée sur le marché français.

2.2 TRIANGLE DE LIQUIDATION

Avant de décrire la méthode de provisionnement de *Chain-Ladder*, il est nécessaire de présenter les triangles de liquidation. En assurance non-vie, les sinistres sont représentés d'une manière agrégée qui s'avère utile et intuitive. En effet, les montants de règlements de sinistres (payés) ou de charges dossier-dossier sont agrégés par année de survenance puis par années de développement. Notons ainsi :

$$\{C_{i,j} : i = 1, \dots, n; j = 1, \dots, n - i + 1\}, \text{ où :}$$

- n est la longueur d'historique (*i.e.* le nombre d'années disponibles) ;
- Le suffixe i fait référence à la ligne et indique l'année de l'accident dite aussi l'année de survenance des sinistres ;
- Le suffixe j fait référence à la colonne et indique l'année de développement ;

- $C_{i,j}$ est le montant des règlements cumulés des sinistres survenus l'année i au bout de j années de développement (dans ce cas on parle de triangles de paiements) **ou** l'évaluation totale dossier-dossier des sinistres survenus l'année i et vue en année j après sa survenance (dans ce cas on parle de triangle de charges).

FIGURE 2.1 – Triangle de liquidation

	1	2	...	j	...	n-1	n
1	$C_{1,1}$	$C_{1,2}$...	$C_{1,j}$...	$C_{1,n-1}$	$C_{1,n}$
2	$C_{2,1}$	$C_{2,2}$...	$C_{2,j}$...	$C_{2,n-1}$	
...		
i	$C_{i,1}$	$C_{i,2}$...	$C_{i,j}$			
⋮	⋮	⋮	⋮				
n-1	$C_{n-1,1}$	$C_{n-1,2}$					
n	$C_{n,1}$						

Remarque 2.2.1 Notons qu'il est aussi possible de représenter les triangles de liquidation par année de souscription, mais dans le cadre de ce mémoire les données seront représentées par **année de survenance**.

Remarque 2.2.2 Les montants $C_{i,j}$ peuvent être considérés bruts ou nets de réassurance. Lors des applications numériques, il sera donc précisé quel type de données est utilisé.

2.3 MODÈLE *Chain-Ladder*

La méthode dite de *Chain-Ladder* est la plus populaire dans le monde de l'assurance non-vie, notamment en France. Elle permet d'estimer (remplir) le triangle inférieur (cf. Figure 2.1) et de déduire la provision relative à ces sinistres. En pratique, cela revient à définir ce qu'on appelle les facteurs de développement $\{\lambda_j, j = 1, \dots, n-1\}$ par :

$$\hat{\lambda}_j = \frac{\sum_{i=1}^{n-j} C_{i,j+1}}{\sum_{i=1}^{n-j} C_{i,j}}. \tag{2.1}$$

Ces facteurs sont ensuite appliqués aux derniers paiements cumulés (ou charges) afin d'estimer le développement futur :

$$\hat{C}_{i,n+2-i} = C_{i,n+1-i} \hat{\lambda}_{n-i+1},$$

$$\hat{C}_{i,k+1} = \hat{C}_{i,k} \hat{\lambda}_k, k = n-i+2, n-i+4, \dots, n-1.$$

Ainsi, la technique de *Chain-Ladder* dans sa forme la plus simple consiste à obtenir des prévisions de la charge "à l'ultime" qui, dans cet exemple, est interprétée comme la dernière année de développement. Afin d'obtenir la provision pour sinistres à payer, il suffit de soustraire de la charge à l'ultime les paiements

déjà réglés pour chaque année de survenance. Ainsi, le montant de provisions à constituer pour l'année de survenance i , noté \hat{R}_i , est égal à :

$$\hat{R}_i = \hat{C}_{i,n} - P_i, \text{ où}$$

P_i est le montant de paiements cumulés relatif à l'année de survenance i . Dans le cas des triangles de paiements cumulés, nous avons bien $P_i = C_{i,n+1-i}$. Le montant total de provisions pour sinistres à payer vaut donc :

$$\hat{R} = \sum_{i=1}^n \hat{R}_i.$$

Remarque 2.3.1 Par abus de langage, nous allons nous référer à ce montant par l'abréviation PSAP. En effet la provision comptable pour sinistres déjà survenus portant le nom de PSAP inclut en plus des frais de gestion de sinistres.

Remarque 2.3.2 En notant $\lambda_{i,j} = \frac{C_{i,j+1}}{C_{i,j}}$ les facteurs de développement individuel, alors, pour tout $j = 1, \dots, n-1$, le facteur développement de $\hat{\lambda}_j$ est la moyenne pondérée des facteurs de développement individuels $\lambda_{i,j}, i = 1, \dots, n-j+1$ par les charges (ou paiements) $C_{i,j}$:

$$\hat{\lambda}_j = \frac{\sum_{i=1}^{n-j} \lambda_{i,j} C_{i,j}}{\sum_{i=1}^{n-j} C_{i,j}}.$$

Remarque 2.3.3 Plusieurs variantes ont été développées à partir de cette méthode et détaillées dans de nombreux mémoires et articles. L'idée générale est de changer la pondération des facteurs de développement individuel, à titre d'exemples :

- Si une année de survenance i présente une sinistralité atypique, il est possible de l'exclure en considérant nul le poids de $\lambda_{i,j}$ lors du calcul de chaque λ_j ;
- Dans le cas où un seul facteur de développement individuel est atypique (par exemple, les tempêtes de fin décembre 1999 Lothar et Martin ont donné lieu à un développement exceptionnel des sinistres entre la fin d'arrêté 1999 et celle à fin 2000), seul le facteur $\lambda_{i,1}$ est exclu ;
- Dans le cas où l'historique est trop profond et certaines années anciennes ne sont plus représentatives du développement récent, les facteurs de développement $\hat{\lambda}_j$ ne sont calculés qu'en utilisant les dernières diagonales du triangle.

Dans le cadre de ce mémoire seule la méthode classique de *Chain-Ladder* sera utilisée.

Remarque 2.3.4 Cette méthode est applicable à toute sorte de triangles, de charges, de paiements, de nombres, de coûts moyens, etc. Au cours de cette étude, elle sera utilisée pour des triangles de paiements cumulés et des triangles de charges, que ce soit bruts ou nets de réassurance.

Cette technique simple et intuitive, utilisée depuis le début des années 1900, a été réécrite dans un modèle actuariel par Mack [9] (voir aussi England et al. [5]). En termes mathématiques, notons \mathcal{F}_n l'information disponible après n -années. Autrement dit, \mathcal{F}_n est la tribu engendrée par l'ensemble :

$$\{C_{i,j}, \forall 2 \leq i + j \leq n + 1\}.$$

Modèle de Chain-Ladder

Sous les deux hypothèses ci-dessous :

Il existe $\lambda_1, \lambda_2, \dots, \lambda_{n-1}$ des facteurs de développement tel que :

$$\forall i = 1, \dots, n, \forall j = 1, \dots, n - 1 : E[C_{i,j+1} | C_{i,1}, \dots, C_{i,j}] = E[C_{i,j+1} | C_{i,j}] = C_{i,j} \lambda_j \tag{2.2}$$

et

$(C_{i,j})_{j \geq 1}$ sont indépendants entre les différentes années de survenance i , (2.3)

nous avons :

$$E[C_{i,n} | \mathcal{F}_n] = C_{i,n+1-i} \prod_{j=n+1-i}^{n-1} \lambda_j \tag{2.4}$$

De plus, les estimateurs $\hat{\lambda}_1, \dots, \hat{\lambda}_{n-1}$ tels que définis par l'équation 2.1 sont sans biais et non corrélés.

Vérification de l'hypothèse relative aux coefficients de développement

L'hypothèse de linéarité 2.2 s'avère importante lors de l'utilisation de la méthode de *Chain-Ladder*. En effet, la méthode *Chain Ladder* suppose que les facteurs de développement individuels ne varient pas en fonction des années de survenance. Dit autrement, les couples $(C_{i,j}, C_{i,j+1})$ doivent être relativement alignés sur une même droite pour une année de développement donnée. Le tableau 2.2 ci-dessous illustre graphiquement les différents points $(C_{i,j}, C_{i,j+1})$ ainsi que la droite $x \rightarrow \hat{\lambda}_i x$ (sur laquelle se retrouvent les points $(C_{i,j}, \hat{\lambda}_i C_{i,j})$).

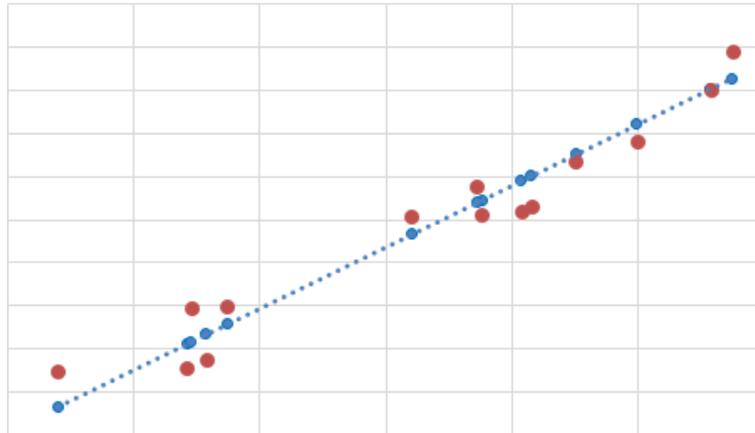


FIGURE 2.2 – Exemple pratique

Afin de vérifier cette hypothèse, nous allons utiliser un indicateur statistique connu sous le nom du coefficient de détermination (dit aussi R^2) entre la série $(C_{i,j+1})_{i=1, n+j-2}$ et celle "estimée" $(\hat{\lambda}_j C_i)_{i=1, n+j-2}$. Nous allons utiliser la notation $R^2(\hat{\lambda}_j)$ afin d'indiquer le coefficient de développement dont la qualité statistique est testée mais aussi pour ne pas confondre avec le montant R qui représente la provision pour sinistres à payer. Pour faciliter les notations, notons $X = (C_{i,j+1})_{i=1, n+j-2}$ et $Y = (\hat{\lambda}_j C_i)_{i=1, n+j-2}$:

$$R2(\hat{\lambda}_j) = \frac{COV(X, Y)^2}{Var(X)Var(Y)},$$

où

- $COV(X, Y)$ désigne la covariance entre les variable X et Y ;
- $Var(X)$ désigne la variance de la variable X .

En pratique, nous estimons les deux quantités $COV(X, Y)$ et $Var(X)$:

- $COV(X, Y) = \frac{1}{N} \sum_{k=1}^N (x_i - \bar{X})(y_i - \bar{Y})$ où N est la longueur des vecteurs X et Y , \bar{X} la moyenne du vecteur X et \bar{Y} la moyenne du vecteur Y ;
- $Var(X) = \frac{1}{N} \sum_{k=1}^N (x_i - \bar{X})^2$.

Remarque 2.3.4 Cet indicateur statistique permet, *in fine*, de donner une indication quantitative, pour un indice j donné, des écarts entre les coefficients de développement individuels $\lambda_{i,j}$ et leur moyenne pondérée λ_j . Ces écarts sont aussi à la base du calcul de l'erreur de prédiction de la méthode de *Chain-Ladder* (cf. section suivante).

2.4 MODÈLE DE MACK

D'un point de vue statistique, étant donné une méthode d'estimation, la prochaine étape consiste à calculer son erreur de prédiction définie comme l'écart-type de la distribution des résultats possibles des réserves ou encore la racine-carrée de l'erreur quadratique (*Mean Square Error of Prediction*) des réserves :

$$MSEP(\hat{R}) = \mathbb{E}[(R - \hat{R})^2].$$

Le modèle actuariel le plus connu est celui de Mack, d'après le nom de Thomas Mack, développé en 1993 dans l'article *Distribution-free calculation of the standard error of Chain Ladder reserves estimates* [9]. Ce modèle est basé sur les hypothèses de travail du modèle de *Chain-Ladder* auxquelles nous ajoutons une hypothèse complémentaire sur l'existence des moments d'ordre 2 afin d'estimer l'erreur de prédiction des réserves.

Modèle de Mack

Si en plus des hypothèses sous-jacentes au modèle de *Chain-Ladder*, il existe $\sigma_1, \sigma_2, \dots, \sigma_{n-1}$ des paramètres de variance tel que :

$$\forall i = 1, \dots, n, \forall j = 2, \dots, n : Var[C_{i,j}|C_{i,1}, \dots, C_{i,j-1}] = Var[C_{i,j}|C_{i,j-1}] = C_{i,j-1}\sigma_j^2 \quad (2.5)$$

alors l'erreur de prédiction des réserves \hat{R}_i par année de survenance peut-être estimée par :

$$M\hat{S}EP(\hat{R}_i) = C_{i,n}^2 \sum_{k=n+1-i}^{n-1} \frac{\hat{\sigma}_k^2}{\hat{\lambda}_k^2} \left(\frac{1}{\hat{C}_{i,k}} + \frac{1}{\sum_{j=1}^{n-k} C_{j,k}} \right) \quad (2.6)$$

où

$$\forall k = 1, \dots, n-2, \hat{\sigma}_k^2 = \frac{1}{n-k-1} \sum_{i=1}^{n-k} C_{i,k-1} \left(\frac{C_{i,k}}{C_{i,k-1}} - \hat{\lambda}_k \right)^2$$

et

$$\hat{\sigma}_{n-1}^2 = \min\left(\frac{\hat{\sigma}_{n-2}^4}{\hat{\sigma}_{n-3}^2}, \min(\hat{\sigma}_{n-3}^2, \hat{\sigma}_{n-2}^2)\right).$$

De plus, les estimateurs $\hat{\sigma}_k^2$ sont sans biais pour $k = 1, \dots, n-2$.

Sous les mêmes hypothèses que le modèle de Mack, l'erreur de prédiction de réserves totales $MSEP(\hat{R})$ peut-être estimée par :

$$M\hat{S}EP(\hat{R}) = \sum_{i=2}^n (M\hat{S}EP(\hat{R}_i) + \hat{C}_{i,n} \left(\sum_{j=i+1}^n \hat{C}_{j,n} \right) \sum_{k=n+1-i}^{n-1} \frac{2\hat{\sigma}_k^2}{\hat{\lambda}_k^2 \sum_{l=1}^{n-1} C_{l,k}}) \quad (2.7)$$

Remarque 2.4.1 Comme indiqué précédemment, les écarts entre les coefficients de développement individuels $\lambda_{i,j}$ et leur moyenne pondérée λ_j sont à la base du calcul de l'erreur de prédiction. En effet, les “ σ de Mack” (langage utilisé pour désigner les montants $\hat{\sigma}_k$) sont la somme de ces écarts $\lambda_{i,j} - \lambda_j$ au carré et pondérés par les montants $C_{i,j}$. Une dégradation des coefficients de détermination $R^2(\zeta_j)$ se traduit par une augmentation des “ σ de Mack” et donc par une augmentation de l'erreur de prédiction. Les indicateurs statistiques $R^2(\zeta_j)$ jouent ainsi un rôle central car ils permettent d'identifier les données de mauvaise qualité.

2.5 MODÈLE DE MERZ & WÜTHRICH

Le modèle de Mack permet d'estimer une erreur de prédiction dite souvent “à l'ultime” car il étudie l'incertitude totale de l'évolution des réserves sur toute la durée de développement des sinistres. Cette vision classique de l'erreur d'estimation de réserves est importante pour la compréhension du risque par les sociétés d'assurance. De plus, cette vision du risque a été également utilisée pour des raisons prudentielles, par exemple en Angleterre le régime prudentiel avant Solvabilité 2, connu sous le nom *Individual Adequacy Capital*, était basé sur une vision du risque “à l'ultime”.

En revanche, la directive Solvabilité 2 a pour objectif de garantir la continuité d'exploitation de la compagnie d'assurance à un horizon d'un an. Autrement dit, la vision du risque est sur un horizon d'un an. Cette nouvelle vision a bousculé les modèles déjà établis puisque jusqu'à récemment, les assureurs analysaient le risque “à l'ultime”. La difficulté de cette méthode provient du fait que la PSAP (ou BE de réserve) est bien calculée “à l'ultime” mais le risque est calculé à un an, dit autrement : **comment la charge à l'ultime peut évoluer à horizon d'un an ?**

Un des modèles les plus connus est celui dit de Merz & Wüthrich. Ce modèle a été introduit en 2008 par Mario Wüthrich, Michael Merz et Natalia Lysenko dans un article intitulé *Uncertainty of the Claims Development Result in the Chain Ladder Method* [11]. Cet article propose une formule fermée, similaire au modèle de Mack, mais qui cette fois permet d'estimer le risque de réserve à un horizon d'un an. Avant d'énoncer le résultat, il est nécessaire d'introduire quelques notations.

Rappelons que l'information (données) disponible après n -années est la tribu \mathcal{F}_n et notons \mathcal{F}_{n+1} l'information disponible la période suivante :

$$\mathcal{F}_{n+1} = \{C_{i,j}, i + j \leq n + 1\} = \mathcal{F}_n \cup \{C_{i,n-i+2}, i = 2, \dots, n\}.$$

L'idée de Merz & Wüthrich est de "recalculer" la PSAP (ou la meilleure estimation) une année plus tard, à date $n + 1$, et de regarder l'écart avec le calcul en n . Pour cela, les auteurs introduisent la notion de **résultat du développement des sinistres** (ou *claim development result*) par année de survénance :

$$\begin{aligned} CDR_i(n+1) &= \mathbb{E}[C_{i,n} | \mathcal{F}_n] - \mathbb{E}[C_{i,n} | \mathcal{F}_{n+1}] \\ &= \mathbb{E}[R_i | \mathcal{F}_n] - \left(X_{i,n+1-i} + \mathbb{E}[R_i^{n+1} | \mathcal{F}_{n+1}] \right) \end{aligned}$$

avec

- $X_{i,n+1-i} = C_{i,n+1-i} - C_{i,n-i}$ et
- $R_i^{n+1} = C_{i,n} - C_{i,n-i}$.

Le résultat global du développement des réserves est la somme des CDR_i par années de survénance :

$$CDR = \sum_{i=2}^n CDR_i.$$

Ce développement des sinistres est appelé par les auteurs "développement **réel**" et qui peut également être interprété comme un résultat théorique (non-observable). Son approximation est appelée par les auteurs "développement observable" $\hat{C}\hat{D}R_i(n+1)$:

$$\hat{C}\hat{D}R_i(n+1) = \hat{C}_{i,n} - \hat{C}_{i,n}^{n+1} = \hat{R}_i - \left(X_{i,n+1-i} + \hat{R}_i^{n+1} \right),$$

avec

- $\hat{C}_{i,n}^{n+1} = C_{i,n+2-i} \prod_{k=n+2-i}^{n-1} \hat{\lambda}_k$,
- $\hat{R}_i = \hat{C}_i^n - C_{i,n-1}$ un estimateur sans biais de $\mathbb{E}[R_i | \mathcal{F}_n]$ et
- $\hat{R}_i^{n+1} = \hat{C}_{i,n}^{n+1} - C_{i,n+1-i}$ un estimateur sans biais de $\mathbb{E}[R_i^{n+1} | \mathcal{F}_{n+1}]$.

Ce "développement observable" peut être interprété comme le résultat enregistré dans le compte de résultat à la fin d'arrêté. C'est en quelque sorte, en termes statistiques, une "réalisation du développement **réel**".

Le modèle de Merz & Wüthrich est basé sur les mêmes hypothèses que celui de Mack avec l'hypothèse supplémentaire que les $(C_{i,j})_{j \geq 1}$ sont une chaîne de Markov. Cette hypothèse est nécessaire afin de s'assurer que l'espérance du résultat du développement des sinistres est nulle (par année de survénance et par conséquence au global) :

$$\mathbb{E}[CDR_i] = 0, \forall i = 2, \dots, n.$$

Cela voudrait dire que pour des facteurs de développement connus λ_i , l'espérance du "développement **réel**" est nulle. En effet, on souhaite que si tout se passe comme prévu alors la charge à l'ultime reste identique pour chaque année de survénance.

Pour revenir au référentiel Solvabilité 2 et au risque de réserve à horizon un an, c'est l'erreur de prédiction du CDR , autour de 0, qu'il serait important d'estimer. Sous l'hypothèse de chaîne de Markov *supra*, l'incertitude de la prédiction du CDR autour de 0 s'exprime simplement par :

$$MSEP(CDR_i) = Var(CDR_i | \mathcal{F}_n) = \mathbb{E}[C_{i,n} | \mathcal{F}_n]^2 \frac{\sigma_{n+1-i}^2}{\lambda_{n+1-i}^2 C_{i,n+1-i}}.$$

Comme les facteurs de développement ne sont pas parfaitement connus, le "développement réel" CDR_i n'est pas observable. L'objectif est donc de calculer :

$$MSEP(C\hat{D}R) = [(C\hat{D}R - 0)^2 | \mathcal{F}_n].$$

Modèle de Merz & Wüthrich

Si en plus des hypothèses du modèle de Mack, les $(C_{i,j})_{j \geq 1}$ sont une chaîne de Markov, alors l'erreur conditionnelle de prédiction autour de 0 du CDR observable est donnée par :

$$M\hat{S}EP_{C\hat{D}R}(0) = \sum_{i=2}^n \hat{C}_{i,n}^2 \hat{\Delta}_i + 2 \sum_{1 < k < i} \hat{C}_{i,n} \hat{C}_{k,n} \hat{\Delta}_k + \sum_{i=2}^n \hat{\Gamma}_i + 2 \sum_{1 < k < i} \hat{Y}_{i,k} \quad (2.8)$$

avec

$$\hat{\Delta}_i = \frac{\hat{\sigma}_{n+1-i}^2}{\hat{\lambda}_{n+1-i}^2 S_{n-i+1}^n} + \sum_{j=n+2-i}^{n-1} \frac{\hat{\sigma}_j^2 C_{n+1-j,j}^2}{\hat{\lambda}_j^2 S_j^n (S_j^{n+1})^2},$$

où $S_u^v = \sum_{k=1}^{v-u} C_{k,u}$, $\forall v > u$;

$$\hat{\Gamma}_i = \hat{C}_{i,n}^2 \left\{ \left(1 + \frac{\hat{\sigma}_{n+1-i}^2}{C_{i,n+1-i} \hat{\lambda}_{n+1-i}^2} \right) \prod_{l=n+2-i}^{n-1} \left(1 + \frac{\hat{\sigma}_l^2 C_{n+1-l}}{\hat{\lambda}_l^2 (S_l^{n+1})^2} \right) - 1 \right\};$$

Et pour $1 < k < i$:

$$\hat{Y}_{i,k} = \hat{C}_{i,n} \hat{C}_{k,n} \left\{ \left(1 + \frac{\hat{\sigma}_{n+1-k}^2}{C_{k,n+1-k} \hat{\lambda}_{n+1-k}^2} \right) \prod_{l=n+2-k}^{n-1} \left(1 + \frac{\hat{\sigma}_l^2 C_{n+1-l}}{\hat{\lambda}_l^2 (S_l^{n+1})^2} \right) - 1 \right\}.$$

2.6 MÉTHODE BOOTSTRAP

Principes généraux

Le *bootstrap* est une technique de ré-échantillonnage (voir [1] pour plus de détail sur la théorie et l'article d'England et al. [4] pour son application au calcul du risque de réserve en assurance non-vie) permettant d'obtenir à partir d'un unique échantillon de données plusieurs informations statistiques généralement obtenues par des outils analytiques. Dans le cadre de l'estimation du risque de réserve, l'idée est de partir du triangle de liquidation et d'une méthode de provisionnement, de reconstruire grâce au ré-échantillonnage un ensemble de "pseudo triangles" et d'appliquer la méthode de provisionnement sous-jacente afin d'obtenir plusieurs réalisations (et donc une distribution) du montant des réserves.

Le principe du *bootstrap* peut s'utiliser pour différentes méthodes statistiques mais aussi pour différents horizons temporels. Dans ce mémoire, le *bootstrap* sera utilisé pour des variations de la technique de Chain-Ladder et pour estimer le risque à un an. Un facteur de queue peut être pris en compte dans la méthode de provisionnement. Notons que la méthode de Merz & Wüthrich ne peut pas tenir compte d'un facteur de queue, la variance étant une formule fermée en fonction des données initiales, c'est-à-dire les $C_{i,j}$. Dans ce mémoire, il ne sera pas tenu compte d'un éventuel facteur de queue afin de pouvoir comparer le résultat du modèle de Merz & Wüthrich avec la méthode de *bootstrap*. En revanche, nous utiliserons des triangles avec une longueur d'historique adéquate pour le risque étudié.

Application de la méthode *Bootstrap* pour l'estimation du risque de réserve à horizon d'un an

Etape 1

Comme indiqué *supra*, la première étape consiste à l'estimation des réserves sur le triangle de liquidation avec la méthode *Chain-Ladder*. Cela permet de calculer les résidus **ajustés** des facteurs de développement individuels conformément à la formule ci-dessous (les résidus sont ajustés de manière à les rendre centrés et réduits) :

$$\forall i, j = 1, \dots, n-1, 2 \leq i+j \leq n, r_{i,j} = \sqrt{\frac{n-j-1}{n-j}} C_{i,j} \frac{\lambda_{i,j} - \hat{\lambda}_j}{\hat{\sigma}_j}.$$

Etape 2

Cette étape, ainsi que celles d'après, se répète plusieurs fois (soit N le nombre de répétitions).

Soit $k \in \llbracket 1, N \rrbracket$. Pour chaque itération k , les résidus $r_{i,j}^{(k)}$ sont tirés de manière aléatoire à l'aide d'une loi uniforme (avec remise) parmi l'ensemble $r_{i,j}, \forall i, j = 1, \dots, n-1, 2 \leq i+j \leq n$. Cela permet de réévaluer les coefficients de passages individuels en inversant la formule *supra* :

$$\forall 1 \leq i+j \leq n, \lambda_{i,j}^{(k)} = \hat{\lambda}_j + r_{i,j}^{(k)} \sqrt{\frac{n-j}{n-j-1}} \sqrt{\frac{\hat{\sigma}_j^2}{C_{i,j}}}.$$

Cela permet de reconstruire le "pseudo-triangle" k par une récurrence "à l'envers" en partant de la dernière diagonale :

$$\forall i = 1, \dots, n, C_{i,n+1-i}^{(k)} = C_{i,n+1-i}$$

et

$$C_{i,j}^{(k)} = \frac{1}{\lambda_{i,j}^{(k)}} C_{i,j+1}^{(k)}, \forall i = 1, \dots, n-1 \text{ et } \forall j = 1, \dots, n-i.$$

Etape 3

L'étape suivante consiste à simuler une nouvelle année ou, dit autrement, une nouvelle diagonale, $C_{i,n+2-i}, \forall i = 2, \dots, n$. Plusieurs choix sont possibles : utiliser les mêmes résidus ou utiliser une loi paramétrique de moyenne $C_{i,n+1-i} \lambda_{n+1-i}^{(k)}$

et de variance $C_{i,n+1-1}\hat{\sigma}_{n-i+1}^2$. Nous allons tester plusieurs choix afin de comparer la matérialité.

Une fois la nouvelle diagonale simulée, des nouveaux coefficients de passage sont calculés en tenant donc compte du “pseudo-triangle” construit à l'étape 2 :

$$\forall j = 2, \dots, n-1, \lambda_j^{(k,n+1)} = \frac{\sum_{i=1}^{n-j+1} C_{i,j+1}^{(k)}}{\sum_{i=1}^{n-j+1} C_{i,j}^{(k)}}.$$

Ces nouveaux coefficients sont utilisés pour la projection des réserves $R^{(k,n+1)}$ par la simple application de la méthode de *Chain-Ladder* au “pseudo-triangle” auquel une nouvelle diagonale a été rajoutée.

2.7 LEMME POUR LA RÉASSURANCE PROPORTIONNELLE

Introduisons d'abord le coefficient de variation ou écart-type relatif. C'est une mesure de dispersion relative d'une distribution égale au rapport entre l'écart-type et sa moyenne. Cette mesure est utilisée en assurance non-vie pour la distribution des réserves. Soit σ_{Mack} le coefficient de variation sous-jacent à la méthode de Mack par :

$$\sigma_{Mack} = \frac{\sqrt{M\hat{S}EP(\hat{R})}}{\hat{R}}.$$

De la même manière, le coefficient de variation sous-jacent à la méthode de Merz & Wüthrich est défini par :

$$\sigma_{MW} = \frac{\sqrt{M\hat{S}EP(\hat{R})}}{\hat{R}}.$$

Afin de diminuer le risque de ruine, et la charge en capital réglementaire, les assureurs utilisent la réassurance comme technique de réduction du risque. Un dispositif de réassurance utilisé fréquemment est la réassurance dite proportionnelle (dite aussi *quote-part*). Un résultat intéressant est que les coefficients de variation sous-jacents aux méthodes de Mack et de Merz & Wüthrich restent inchangés dans le cas d'une réassurance proportionnelle.

Pour $\alpha > 0$ et pour un triangle de liquidation $(C_{i,j})_{2 \leq i+j \leq n+1}$, appelons le α -triangle, le triangle de liquidation $(\alpha C_{i,j})_{2 \leq i+j \leq n+1}$. Sous les hypothèses du modèle de Mack, les coefficients de variation sous-jacents à la méthode de Mack pour ces deux triangles sont identiques.

Démonstration. Nous observons aisément que les coefficients de développement des deux triangles sont identiques, $\hat{\lambda}_j = \hat{\lambda}_j^{(\alpha)}, \forall j = 1, \dots, n-1$. Cela permet de déduire facilement que $\hat{R}^{(\alpha)} = \alpha \hat{R}$.

Calculons les “ σ de Mack” du α -triangle en fonction du triangle d'origine. Soit $k = 1, \dots, n-2$:

$$(\hat{\sigma}_k^{(\alpha)})^2 = \frac{1}{n-k-1} \sum_{i=1}^{n-k} \alpha C_{i,k-1} \left(\frac{\alpha C_{i,k}}{\alpha C_{i,k-1}} - \hat{\lambda}_k^{(\alpha)} \right)^2$$

$$= \frac{1}{n-k-1} \sum_{i=1}^{n-k} \alpha C_{i,k-1} \left(\frac{C_{i,k}}{C_{i,k-1}} - \hat{\lambda}_k \right)^2 = \alpha \hat{\sigma}_k^2.$$

La même relation est obtenue pour $k = n - 1$.

Calculons ensuite l'erreur quadratique des réserves $\hat{R}_i^{(\alpha)}$. Soit $k = 1, \dots, n - 2$:

$$\begin{aligned} M\hat{S}EP(\hat{R}_i^{(\alpha)}) &= (\alpha C_{i,n})^2 \sum_{k=n+1-i}^{n-1} \frac{(\hat{\sigma}_k^{(\alpha)})^2}{(\hat{\lambda}_j^{(\alpha)})^2} \left(\frac{1}{\alpha \hat{C}_{i,k}} + \frac{1}{\sum_{j=1}^{n-k} \alpha C_{j,k}} \right) \\ &= \alpha^2 C_{i,n}^2 \sum_{k=n+1-i}^{n-1} \alpha \frac{\hat{\sigma}_k^2}{\hat{\lambda}_k^2} \frac{1}{\alpha} \left(\frac{1}{\hat{C}_{i,k}} + \frac{1}{\sum_{j=1}^{n-k} C_{j,k}} \right) = \alpha^2 M\hat{S}EP(\hat{R}_i). \end{aligned}$$

Afin de montrer que $M\hat{S}EP(\hat{R}^{(\alpha)}) = \alpha^2 M\hat{S}EP(\hat{R})$, il reste à montrer que :

$$\alpha \hat{C}_{i,n} \left(\sum_{j=i+1}^n \alpha \hat{C}_{j,n} \right) \sum_{k=n+1-i}^{n-1} \frac{2(\hat{\sigma}_k^{(\alpha)})^2 (\hat{\lambda}_k^{(\alpha)})^2}{\sum_{l=1}^{n-1} \alpha C_{l,k}} = \alpha^2 \hat{C}_{i,n} \left(\sum_{j=i+1}^n \hat{C}_{j,n} \right) \sum_{k=n+1-i}^{n-1} \frac{2\hat{\sigma}_k^2 \hat{\lambda}_k^2}{\sum_{l=1}^{n-1} C_{l,k}}.$$

Cela s'obtient de la même manière que les relations ci-dessous en utilisant $(\hat{\sigma}_k^{(\alpha)})^2 = \alpha \hat{\sigma}_k^2$ et $\hat{\lambda}_j = \hat{\lambda}_j^{(\alpha)}$, ce qui achève la preuve car nous obtenons bien :

$$\sigma_{Mack}^{(\alpha)} = \frac{\sqrt{M\hat{S}EP(\hat{R}^{(\alpha)})}}{\hat{R}^{(\alpha)}} = \frac{\sqrt{\alpha^2 M\hat{S}EP(\hat{R})}}{\alpha \hat{R}} = \frac{\sqrt{M\hat{S}EP(\hat{R})}}{\hat{R}} = \sigma_{Mack}.$$

□

De la même manière, nous pouvons montrer un résultat similaire pour le coefficient de variation sous-jacent au modèle de Merz & Wüthrich. Pour $\alpha > 0$ et pour un triangle de liquidation $(C_{i,j})_{2 \leq i+j \leq n+1}$, appelons le α -triangle, le triangle de liquidation $(\alpha C_{i,j})_{2 \leq i+j \leq n+1}$. Sous les hypothèses du modèle de Merz & Wüthrich, les coefficients de variation sous-jacents à la méthode de Merz & Wüthrich pour ces deux triangles sont identiques.

Remarquons que ce résultat n'est plus valable si le triangle initial n'est plus perturbé par une fonction linéaire de type $x \rightarrow \alpha x$ mais par une fonction affine $x \rightarrow \alpha x + \beta$ avec $\beta \neq 0$. Une démonstration empirique sera donnée ci-dessous, mais une simple intuition est suffisante. En effet, une transformation affine $x \rightarrow \alpha x + \beta$ change les coefficients de développement individuels $\lambda_{i,j}$ ainsi que les coefficients de développement λ_j .

À des fins de comparaison (et de démonstration empirique), nous indiquerons ci-dessous le changement du coefficient de variation d'un triangle perturbé par une fonction affine $x \rightarrow x + \beta$. Comme une transformation linéaire ne change pas le coefficient de variation, le triangle initial sera transformé dans un α -triangle de manière à avoir $C_{n,1} = 1000$. Considérons le triangle de base ci-dessous (tableau 2.3).

Le tableau 2.4 donne les évolutions du coefficient de variation (de Merz & Wüthrich et de Mack) quand le triangle de base est perturbé par une transformation $x \rightarrow x + \beta$ avec $\beta = 10^k$, $k = -1, \dots, 7$.

FIGURE 2.3 – Triangle de paiements cumulés

	1	2	3	4	5	6	7	8	9	10
1	749	1004	1027	1039	1048	1075	1087	1094	1105	1106
2	822	1119	1148	1161	1186	1208	1232	1239	1240	
3	847	1184	1220	1260	1281	1303	1317	1317		
4	825	1090	1143	1172	1198	1228	1229			
5	806	1122	1174	1204	1232	1236				
6	920	1223	1273	1303	1311					
7	870	1167	1213	1225						
8	900	1233	1253							
9	988	1264								
10	1000									

FIGURE 2.4 – Résultats de la perturbation par une fonction affine $x \rightarrow x + \beta$

β	Moyenne Chain-Ladder	$\sqrt{MSEP_{MW}}$	σ_{MW}	$\sqrt{MSEP_{Mack}}$	σ_{Mack}
$\beta = 0$	907,7	79,1	8,71%	110,5	12,18%
$\beta = 10^{-1}$	907,7	79,1	8,71%	110,5	12,18%
$\beta = 10^0$	907,6	79,1	8,71%	110,5	12,18%
$\beta = 10^1$	906,6	78,9	8,71%	110,3	12,17%
$\beta = 10^2$	897,1	77,1	8,60%	108,5	12,09%
$\beta = 10^3$	851,7	70,0	8,23%	100,4	11,79%
$\beta = 10^4$	807,7	64,8	8,03%	93,6	11,59%
$\beta = 10^5$	798,5	63,9	8,01%	92,3	11,56%
$\beta = 10^6$	797,6	63,8	8,00%	92,2	11,56%

Conclusion Ce résultat est important pour ce mémoire car il permettra **d'anonymiser les triangles**. Perturber un triangle par une transformation linéaire $x \rightarrow \alpha x$ change le triangle mais garde inchangés les coefficients de développement individuels ce qui permettrait éventuellement l'identification de données (même si cela ne serait possible que par la société d'où proviennent les données). En revanche, une perturbation de type $x \rightarrow \alpha x + \beta$ avec $\beta \neq 0$ rend impossible l'identification des données utilisées. De plus, comme le montre le tableau précédent, **les coefficients de variation changent peu** si β est de l'ordre de $C_{i,j}$.

2.8 LIEN AVEC LA CHARGE EN CAPITAL RÉGLEMENTAIRE SOLVABILITÉ 2 EN FORMULE STANDARD

En formule standard, l'exigence en capital pour le risque de réserve est intégrée dans le sous-module de "risque de primes et de réserve" qui est défini (Article 105.2 de la Directive Solvabilité 2 - voir [6]) comme "le risque de perte, ou de changement défavorable de la valeur des engagements d'assurance, résultant de fluctuations affectant la date de survenance, la fréquence et la gravité des événements assurés, ainsi que la date et le montant des règlements de sinistres". L'article 115 du règlement délégué [7] définit la charge en capital pour le risque de primes et de réserve en non-vie comme :

$$SCR_{nl\text{premres}} = 3\sigma_{nl}V_{nl}$$

où V_{nl} et σ_{nl} sont définis aux articles 116 et respectivement 117 du règlement délégué [7].

La mesure de volume V_{nl} est définie comme “la somme des mesures du volume pour le risque de primes et de réserve des segments visés à l’annexe II” (i.e. pour les différentes lignes d’activité) :

$$V_{nl} = \sum_s V_s$$

avec

$$V_s = (V_{s,prem} + V_{s,res})(0,75 + 0,25DIV_s).$$

DIV_s est un facteur qui tient compte de la diversification géographique mais qui est égal à 1 pour une société d’assurance ayant une activité dans un unique pays. En pratique, la mesure de volume pour une ligne d’activité s est la simple somme entre les mesures de volumes pour le risque de primes et pour le risque de réserve :

$$V_s = (V_{s,prem} + V_{s,res}).$$

Pour le risque de réserve, la mesure de volume est la meilleure estimation de la provision pour sinistres (“*BE de réserve*”) nette de réassurance. La mesure de volume pour le risque de primes est définie à l’article 116.3 mais n’est pas développée davantage à cause de son inutilité dans le cadre de ce mémoire.

L’écart-type du risque de primes et de réserve est défini à l’article 117 du règlement délégué [7] par :

$$\sigma_{nl} = \frac{1}{V_{nl}} \sqrt{\sum_{s,t} Corr(s,t) \sigma_s V_s \sigma_t V_t}$$

où :

- $Corr(s,t)$ est une matrice de corrélation définie à l’annexe IV du règlement délégué.
- V_s et V_t sont des mesures de volumes des segments s et t
- σ_s et σ_t sont “les écarts types du risque de primes et de réserve en non-vie des segments s et t , respectivement” ;
- la définition de σ_s est donnée par :

$$\sigma_s = \frac{\sqrt{\sigma_{s,prem}^2 V_{s,prem}^2 + \sigma_{s,res}^2 V_{s,res}^2 + \sigma_{s,prem} V_{s,prem} \sigma_{s,res} V_{s,res}}}{V_s},$$

où $\sigma_{s,prem}$ et $\sigma_{s,res}$ sont “l’écart type du risque de primes en non-vie” et respectivement “l’écart type du risque de réserve en non-vie” du segment s , tels que définis à l’annexe II du règlement délégué.

Remarque 2.8.1 Le règlement délégué appelle $\sigma_{s,prem}$, $\sigma_{s,res}$ et σ_s des écarts-types, mais c’est un abus de langage. Ces valeurs sont exprimées en pourcentage et correspondent à ce qu’on appelle dans ce mémoire des coefficients de variations (ou CoV en abrégé comme *coefficient of variation*).

Rappelons ci-dessous les coefficients de variation des segments d’activités (Solvabilité 2) les plus importantes et qui seront étudiées dans le cadre de ce

Ligne d'activité	CoV risque de Primes	CoV risque de réserve
Frais médicaux	5%	5,7%
Protection du revenu	8,5%	14%
Responsabilité civile automobile	10%	9%
Autre assurance des véhicules à moteur	8%	8%
Incendie et autres dommages aux biens	8%	10%
Responsabilité civile générale	14%	11%
Protection juridique	8,3%	5,5%
Assistance	6,4%	22%
Pertes pécuniaires diverses	13%	20%

FIGURE 2.5 – CoV par ligne d'activité

mémoire.

Remarque 2.8.2 Pour les segments d'activité de responsabilité civile automobile, de responsabilité civile générale et d'assurance incendie et autres dommages aux biens, l'article 117.3 du règlement délégué précise que "l'écart type du risque de primes en non-vie d'un segment donné est égal au produit de l'écart type du risque de primes brut en non-vie et du facteur d'ajustement pour la réassurance non proportionnelle [...] est égale à 80%". Comme la grande majorité des assureurs se protègent contre le risque de pointe sur ces lignes d'activité, le consensus de marché est de calculer l'exigence en capital pour le risque de primes en appliquant cet abattement ce qui revient à considérer les CoV égaux à 8%, 6,4% et respectivement 11,2%. Cela n'a pas de conséquence sur ce mémoire qui se concentre sur le risque de réserve.

Remarque 2.8.3 Pour les lignes d'activité de frais médicaux, de protection juridique et d'assistance, les CoV ci-dessus sont ceux après la modification du règlement délégué [8] qui a eu lieu en 2019. Pour rappel, les CoV du risque de réserves sur ces lignes d'activité au moment de la rentrée en vigueur du régime prudentiel Solvabilité 2 début 2016 étaient de respectivement, 5%, 20% et 20%. Ces évolutions seront commentées à la lumière des résultats dans les prochains chapitres de ce mémoire.

En revenant à la charge en capital au titre du risque de réserve, celle-ci pourrait être calculée en prenant une société d'assurance fictive avec une seule ligne d'activité s et en supposant qu'il n'y ait pas de nouvelles primes. Ainsi, le risque de primes et réserve en non-vie reviendrait juste à un risque de réserve dont la charge en capitale serait :

$$SCR_s = 3\sigma_{s,res}V_{s,res}.$$

Pour rappel, l'article 101 de la directive Solvabilité 2 (voir [6]) définit le capital de solvabilité requis comme "la valeur en risque des fonds propres de base de l'entreprise d'assurance ou de réassurance, avec un niveau de confiance de 99,5% à horizon d'un an". Dit autrement, étant donné la distribution de pertes liée à un risque, l'exigence en capital est la différence entre le quantile à 99,5% et la moyenne. Prenons X une variable aléatoire, alors :

$$SCR = q_{99,5\%}(X) - E[X].$$

Nous pouvons réécrire cette formule différemment :

$$SCR = q_{99,5\%} \left(E[X] \frac{X}{E[X]} \right) - E[X] = E[X] q_{99,5\%} \left(\frac{X}{E[X]} \right) - E[X] = E[X] \left(q_{99,5\%} \left(\frac{X}{E[X]} \right) - 1 \right).$$

Notons le risque “normalisé” par $Z = \frac{X}{E[X]}$. Le lien entre les coefficients de variation de X et Z est donné par la formule ci-dessous car l’espérance de Z vaut 1 :

$$CoV(X)^2 = Var(Z).$$

L’hypothèse centrale de la formule standard est que le risque suit une **loi log-normale**. Soient μ_Z et σ_Z les paramètres de la loi log-normale Z . Comme

$$E[Z] = 1 = \exp\left(\mu_Z + \frac{1}{2}\sigma_Z^2\right),$$

nous obtenons :

$$\mu_Z = -\frac{1}{2}\sigma_Z^2.$$

La variance d’une loi log-normale de paramètres μ_Z et σ_Z vaut :

$$Var[Z] = \left(\exp(\sigma_Z^2) - 1\right) \exp(2\mu_Z + \sigma_Z^2).$$

Ce qui dans le cas du risque “normalisé” se réduit à :

$$Var[Z] = \exp(\sigma_Z^2) - 1.$$

Par ailleurs nous avons :

$$q_{99,5\%}(Z) = \exp(\sigma_Z \Phi_{99,5\%}^{-1} + \mu_Z),$$

où $\Phi_{99,5\%}^{-1}$ est le quantile à 99,5% d’une loi normale standard (centrée et réduite). Nous obtenons ainsi :

$$q_{99,5\%}(Z) - 1 = \frac{\exp(\Phi_{99,5\%}^{-1} \sqrt{\ln(CoV(X)^2 + 1)})}{\sqrt{CoV(X)^2 + 1}} - 1.$$

Un développement limité en 0 permet de montrer que :

$$\frac{\exp(\Phi_{99,5\%}^{-1} \sqrt{\ln(CoV(X)^2 + 1)})}{\sqrt{CoV(X)^2 + 1}} - 1 \sim \Phi_{99,5\%}^{-1} CoV(X).$$

Donc nous obtenons :

$$SCR \approx \Phi_{99,5\%}^{-1} CoV(X) E[X].$$

Comme, $\Phi_{99,5\%}^{-1} \approx 2,58$, nous comprenons donc l’approximation $SCR_s = 3\sigma_{s,res} V_{s,res}$.

Remarque 2.8.4 Notons également que pour les valeurs de CoV entre 5% et 22% (voir tableau 2.5 ci-dessus) :

$$\frac{SCR}{CoV \times E[X]} \in [2,72; 3,23].$$

La formule “ $3\sigma V$ ” est précise pour un coefficient de variation $CoV \approx 14,5\%$. À l’exception des lignes d’activité d’assistance et de pertes pécuniaires diverses,

la plupart des coefficients de variation sont autour de 10% et pour cette valeur $\frac{SCR}{CoV \times E[X]} \approx 2,87$. En appliquant la formule “ $3\sigma V$ ”, l'exigence en capital réglementaire en formule standard surestime d'environ 5% la charge en capital au titre du risque de primes et de réserves en non-vie par rapport à un calcul exact. Cet écart peut être cependant interprété comme une marge de prudence pour tenir compte de l'erreur de modèle (i.e. l'hypothèse de log-normalité de l'ensemble des distributions de risques pour chaque ligne d'activité que ce soit pour le risque de réserve ou pour le risque de primes).

2.9 MÉTHODE RÉGLEMENTAIRE POUR LES PARAMÈTRES SPÉCIFIQUES

Ce que les textes stipulent

Pour le calcul de leur exigence de capital réglementaire, les organismes ou groupes d'assurance peuvent sous certaines conditions utiliser des paramètres qui leur sont propres en remplacement des paramètres de la formule standard. Ces paramètres sont appelés couramment USP (*Undertaking specific parameters*), ou bien GSP (*Group specific parameters*) quand ils sont appliqués au niveau d'un groupe pour le calcul de l'exigence de capital sur base combinée ou consolidée. En effet, l'article R352-5 du Code des Assurances [3] précise que : “Sous réserve de l'accord de l'Autorité de contrôle prudentiel et de résolution, les entreprises d'assurance et de réassurance peuvent, lorsqu'elles calculent les modules risque de souscription en vie, risque de souscription en non-vie et risque de souscription en santé remplacer, dans la conception de la formule standard, un sous-ensemble des paramètres précisés à l'article 218 du règlement délégué [...] par des paramètres qui sont propres à l'entreprise concernée. Ces paramètres sont calibrés sur la base des données internes de l'entreprise concernée ou de données directement pertinentes pour les opérations de cette entreprise, sur la base de méthodes standardisées.”

L'article 218 du règlement délégué [7] précise que les paramètres pouvant potentiellement être remplacés par des USP pour tous les segments non-vie et santé NSLT (non similaire à la vie) sont :

- soit le coefficient de variation brut de réassurance non-proportionnelle du risque de primes (une méthode disponible), soit le facteur d'ajustement lié à la réassurance non-proportionnelle du risque de primes, si pour le segment considéré il existe un contrat de réassurance en excédent de sinistres éligible ;
- le coefficient de variation net de réassurance non-proportionnelle du risque de primes ;
- le coefficient de variation du risque de réserve.

Les méthodes statistiques relatives à l'évaluation des paramètres spécifiques sont entièrement standardisées, mais leur utilisation est conditionnée à la qualité des données et au profil de risque de l'organisme. En ce qui concerne le risque de réserve, l'article 220 du règlement délégué [7] préconise deux méthodes décrites à l'annexe XVII du règlement délégué :

- La première méthode revient à calibrer une loi-lognormale sur les boni-mali du BE de réserve (par boni-mali on comprend l'écart entre le BE de réserve calculé en N et la somme du BE de réserve en $N+1$ plus les paiements). Cette méthode ne sera pas étudiée dans ce mémoire car seulement 4 ans se sont écoulés depuis l'entrée en vigueur du régime prudentiel Solvabilité 2 et cet historique est assez limité pour considérer les résultats pertinents ;
- La deuxième méthode est très similaire à celle de Merz & Wüthrich et nous allons étudier les différences.

Remarquons d'abord que le calcul final du paramètre propre (USP) par ligne d'activité $\sigma_{res,s,USP}$ est une moyenne pondérée entre l'USP calculé avec la méthode décrite dans l'annexe du règlement délégué et celui de la formule standard :

$$\sigma_{res,s,USP} = c\sigma_{USP,s} + (1 - c)\sigma_{res,s}.$$

Le paramètre c est un paramètre de crédibilité qui dépend de la longueur d'historique utilisée pour calibrer $\sigma_{USP,s}$ (que nous allons noter σ_{USP} afin d'alléger les notations). Comme nous souhaitons comprendre la différence avec le modèle de Merz & Wüthrich, dans ce mémoire nous ne nous intéressons pas au paramètre c mais juste au coefficient de variation σ_{USP} .

Remarque 2.9.1 La méthode réglementaire pour le calcul du coefficient de variation pour le risque de réserve est basée sur des triangles de paiements cumulés : “Les données servant à l'estimation de l'écart type propre à l'entreprise pour le risque de réserve en non-vie ou le risque de réserve en santé non-SLT dans le segment s consistent en montants de paiements cumulés relatifs aux sinistres d'assurance ou de réassurance dans le segment s (montants de sinistres cumulés), indiqués séparément pour chaque année d'accident et chaque année de développement des paiements.”

Remarque 2.9.2 Afin de faciliter la lecture, nous ne gardons pas les notations du règlement délégué et nous utiliserons les notations déjà utilisées précédemment.

Méthodologie de calcul des paramètres propres pour le risque de réserve

Comme indiqué *supra*, la première méthode revient à calibrer une loi lognormale sur le boni-mali du BE de réserve, mais qui n'est pas d'intérêt pour ce mémoire. C'est bien la deuxième méthode qui nous intéresse car en fait elle est une approximation de l'erreur quadratique “originale” de Merz & Wüthrich. En effet, les mêmes auteurs ont publié un deuxième article *Modelling The Claims Development Result For Solvency Purposes* [10] juste après le premier où, en partant d'un développement limité, ils ont simplifié la formule initiale.

Soient $a_i \in \mathbb{R}^+$. Sous l'hypothèse $a_i \ll 1$ nous avons l'approximation :

$$\prod_{i=1}^N (1 + a_i) = 1 + \sum_{i=1}^N a_i.$$

En utilisant ce développement limité, les auteurs simplifient ainsi “les termes” Γ_i et Y_i (voir section 2.5), pour rappel :

$$\hat{\Gamma}_i = \hat{C}_{i,n}^2 \left\{ \left(1 + \frac{\hat{\sigma}_{n+1-i}^2}{C_{i,n+1-i} \hat{\lambda}_{n+1-i}^2} \right) \prod_{l=n+2-i}^{n-1} \left(1 + \frac{\hat{\sigma}_l^2 C_{n+1-l}}{\hat{\lambda}_l^2 (S_l^{n+})^2} \right) - 1 \right\};$$

Et pour $k < i$:

$$\hat{Y}_{i,k} = \hat{C}_{i,n} \hat{C}_{k,n} \left\{ \left(1 + \frac{\hat{\sigma}_{n+1-k}^2}{C_{k,n+1-k} \hat{\lambda}_{n+1-k}^2} \right) \prod_{l=n+2-k}^{n-1} \left(1 + \frac{\hat{\sigma}_l^2 C_{n+1-l}}{\hat{\lambda}_l^2 (S_l^{n+})^2} \right) - 1 \right\}.$$

Afin de simplifier ces expressions en utilisant le développement limité indiqué *supra*, les auteurs font implicitement l'hypothèse supplémentaire suivante :

$$\frac{\hat{\sigma}_j^2}{\hat{\lambda}_j^2 C_j} \ll 1.$$

Modèle simplifié de Merz & Wüthrich

Sous les hypothèses du modèle de Merz & Wüthrich, si en plus les quantités $\frac{\hat{\sigma}_j^2}{\hat{\lambda}_j^2 C_j}$ sont suffisamment petites, alors l'erreur conditionnelle de prédiction autour de 0 du CDR observable peut-être approximé par :

$$M\hat{S}EP_{sCDR}(0) = \sum_{i=2}^n \hat{C}_{i,n}^2 \left(\hat{\Lambda}_i + \hat{\phi}_i + \hat{\Psi}_i \right) + 2 \sum_{1 < k < i} \hat{C}_{i,n} \hat{C}_{k,n} \left(\hat{\Lambda}_k + \hat{\Xi} \right) \quad (2.9)$$

avec

$$\begin{aligned} \hat{\phi}_i &= \sum_{j=n+2-i}^{n-1} \frac{\hat{\sigma}_j^2 C_{n+1-j,j}}{\hat{\lambda}_j^2 S_j^{n+1}}; \\ \hat{\Psi}_i &= \frac{\hat{\sigma}_{n+1-i}^2}{\hat{\lambda}_{n+1-i}^2 C_{i,n+1-i}}; \\ \hat{\Lambda}_i &= \frac{\hat{\sigma}_{n+1-i}^2 C_{i,n+1-i}}{\hat{\lambda}_{n+1-i}^2 S_{n+1-i}^{n+1} S_{n+1-i}^n} + \sum_{j=n+2-i}^{n-1} \frac{\hat{\sigma}_j^2 C_{n+1-j,j}^2}{\hat{\lambda}_j^2 S_j^n (S_j^{n+1})^2}; \\ \hat{\Xi}_i &= \hat{\phi}_i + \frac{\hat{\sigma}_{n+1-i}^2}{\hat{\lambda}_{n+1-i}^2 S_{n+1-i}^{n-1}}. \end{aligned}$$

2.10 CALIBRAGE DE LA FORMULE STANDARD PAR EIOPA

EIOPA a mené plusieurs études d'impact sur le calibrage des coefficients de variation pour les lignes d'activité non-vie et santé non-similaire à la vie. La dernière date de 2011 et elle est connue sous le nom de **QIS5** pour *Quantitative impact study* numéro 5 dont les résultats ont été publiés en décembre 2011 dans le document *Calibration of the Premium and Reserve Risk Factors in the Standard Formula of Solvency II - Report of the Joint Working Group on Non-Life and Health NSLT Calibration* [2]. Dans cette section, nous présentons une synthèse de l'approche scientifique retenue et des résultats mis en valeur par ce groupe de travail. Notons tout d'abord que le choix final des coefficients de variation est le résultat d'une concertation européenne incluant l'ensemble des superviseurs nationaux mais également l'industrie : ce choix ne sera pas commenté pour la suite, seule l'approche académique nous intéresse ici.

Données utilisées

Pour le calibrage du risque de réserve, l'étude a été menée en utilisant les données suivantes :

- les triangles de paiements ;
- des triangles des meilleures estimations de sinistres à payer ;
- des triangles de charges.

De plus, il a été demandé aux sociétés d'assurance de renseigner (pour celles qui étaient en mesure de le faire) les triangles bruts et nets de réassurance incluant ou non les sinistres de catastrophes naturelles (également bruts et nets de réassurance).

Comme le nombre de données nettes de réassurance était bien inférieur au nombre de données brutes de réassurance, EIOPA a construit l'étude sur les triangles bruts de réassurance. Ensuite, EIOPA a appliqué un abattement "gross-to-net". Dans le cadre de ce mémoire, nous ferons des études séparées entre les triangles bruts et nets de réassurance.

Pour le calibrage final, certaines sociétés d'assurance ont été retirées quand les données ont été considérées de mauvaises qualité. EIOPA a regardé la distribution de l'ensemble des coefficients de variations (peu importe la méthode considérée) et a retiré les valeurs extrêmes de manière à diminuer le kurtosis. Nous allons également suivre un algorithme afin de retirer les données de mauvaise qualité et qui sera détaillé dans le prochain chapitre. En revanche, cet algorithme ne sera pas fondé sur des valeurs extrêmes car le but de ce mémoire est d'analyser la dépendance du coefficient de variation par rapport au volume. Alors que la méthodologie EIOPA consistait à enlever les valeurs très faibles du coefficient de variation, nous les garderons si les indicateurs statistiques ne montrent pas des problèmes de qualité de données.

Concernant la longueur de l'historique (i.e. des triangles) utilisés, EIOPA a mené une analyse de sensibilité sur l'impact du filtrage de la base de données pour exclure les ensembles de données à trois niveaux différents :

- Exclure les ensembles de données avec moins de trois ans d'historique de données ;
- Exclure les ensembles de données avec moins de cinq ans d'historique de données ;
- Exclure les ensembles de données avec moins de sept ans d'historique de données.

Les résultats étant sensiblement proches, EIOPA a gardé la première option afin d'avoir un maximum de données. Dans ce mémoire, nous préférons garder des triangles avec un minimum 4 années d'historique. Comme indiqué dans la section 2.6, pour les segments d'activité à développement plus long nous enlèverons les triangles de taille courte.

Méthodes utilisées

Dans le cadre du QIS5, EIOPA a utilisé les mêmes méthodes que celle proposées aujourd'hui dans le règlement délégué pour le calcul des paramètres spécifiques (USP), pour rappel :

- une méthode basée sur les boni-mali du BE de réserves ;
- une méthode proche du modèle de Merz & Wüthrich.

Ensuite, en partant du modèle simplifié expliqué dans l'introduction (i.e. le capital pour éviter le risque de ruine augmente comme la racine carrée du volume), EIOPA a utilisé un modèle de minimisation de l'erreur quadratique pour estimer le coefficient de variation pour chaque ligne d'activité :

$$\sum_c \left(\sigma_{lob} V_{c,lob} - \sqrt{MSEP(R)_{c,lob}} \right)^2,$$

où :

- c est un indice qui prend ses valeurs dans l'ensemble des sociétés de l'échantillon ;
- $V_{c,lob}$ est une mesure de volume et qui a été considérée égale au résultat de l'application de la méthode de Chain-Ladder ;
- $MSEP(R)_{c,lob}$ est une mesure de l'erreur quadratique (en fonction du modèle utilisé).

En pratique, en résolvant un problème de minimum absolu, nous obtenons en dérivant :

$$\hat{\sigma}_{lob} = \frac{\sum_c V_{c,lob} \sqrt{MSEP(R)_{c,lob}}}{\sum_c V_{c,lob}^2} = \frac{\sum_c V_{c,lob}^2 CoV_{c,lob}}{\sum_c V_{c,lob}^2}. \quad (2.10)$$

Quelques conclusions d'EIOPA sur le risque de réserve

En plus de l'estimation et d'un coefficient de variation "moyen" par ligne d'activité en Europe, EIOPA a mis aussi en valeur quelques faits stylisés. Ceux-ci sont représentés par le graphe ci-dessous montrant les valeurs des coefficients de variation par société en fonction de la mesure de volume. Notons tout d'abord qu'EIOPA n'a publié qu'une seule figure de ce type sur la ligne d'activité "responsabilité civile auto". De plus, EIOPA a précisé que ce graphe est "typique" pour l'ensemble des segments d'activité mais sans fournir d'autres exemples ou commentaires.

A partir de ce graphe, deux conclusions ont été développées par EIOPA :

- Le coefficient de variation diminue avec l'augmentation du volume de l'activité ;
- Pour les très grands portefeuilles, il existe également un niveau plancher strictement positif pour la volatilité (i.e. la volatilité ne diminue pas à zéro).

De plus, nous remarquons que ce graphe correspond au risque de primes car le volume représenté est le niveau de primes émises et que tous les points avec des volumes "extrêmes" ont été retirés.

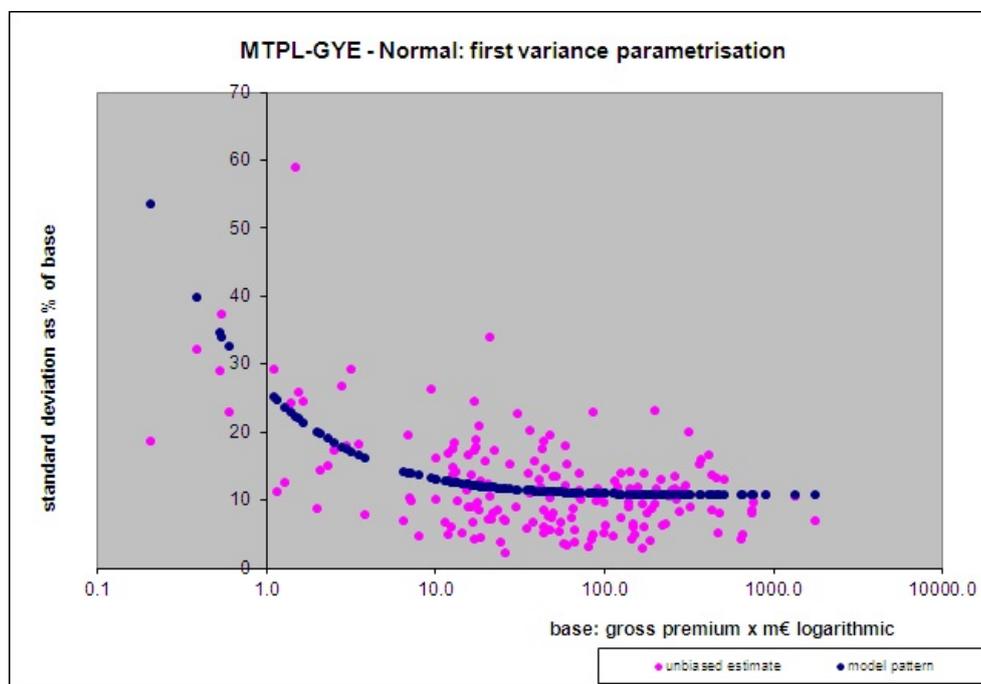


FIGURE 2.6 – Exemple dépendance CoV en fonction de volume - source [2]

Méthodologie de travail et comparaison des méthodes

3.1	Description des données utilisées	30
3.2	Contrôle de la qualité des données	33
3.3	Algorithme de traitement des données	36
3.4	Retraitement des données	37
3.5	Triangles de charges inexploitable	47
3.6	Différence entre les écarts-types des méthodes de Merz & Wüthrich et Bootstrap	48
3.7	Convergence dans la méthode Bootstrap	49
3.8	Méthode Bootstrap avec d'autres lois de simulation	50

3.1 DESCRIPTION DES DONNÉES UTILISÉES

Sources des données utilisées

La réglementation Solvabilité 2 comprend trois piliers : le pilier 1 relatif aux exigences quantitatives, le pilier 2 relatif aux exigences qualitatives et le pilier 3 relatif aux exigences de reporting et d'information. Le pilier 3 comprend notamment un reporting régulier à l'autorité de contrôle sous la forme de rapports narratifs et d'états quantitatifs. Les états quantitatifs Solvabilité 2, dits *Quantitative Report Templates* (ou QRT), sont remis au superviseur dans leur intégralité à fréquence annuelle. Ils comprennent des informations relatives à 8 catégories : le bilan prudentiel, les fonds propres, l'exigence de capital, l'analyse de variation de fonds propres, les placements, les provisions techniques, la réassurance et des états spécifiques pour les groupes. Dans le cadre de ce mémoire, les données utilisées ont été tirées des QRTs solos des sociétés d'assurance, et plus particulièrement des états s.19.01.01. Ces états contiennent, pour chaque société et chaque segment d'activité exercée, l'ensemble des "triangles" de liquidation.

Les segments d'activité Solvabilité 2 considérés

Les segments d'activité Solvabilité 2 étudiés sont détaillés ci-dessous et regroupent l'ensemble des activités relatives aux engagements d'assurance non-vie et santé non-similaire à la vie. Un segment d'activité Solvabilité 2 contient les affaires directes et les engagements de réassurance proportionnelle¹. Nous reprenons ci-dessous leur définition telle que précisée par l'annexe du règlement délégué :

1. **Assurance des frais médicaux (code QRT : x69)**
Engagements d'assurance de frais médicaux dans les cas où l'activité sous-jacente n'est pas exercée sur une base technique similaire à celle de l'assurance vie ;
2. **Assurance de protection du revenu (code QRT : x52)**
Engagements d'assurance de protection du revenu dans les cas où l'activité sous-jacente n'est pas exercée sur une base technique similaire à celle de l'assurance vie ;
3. **Assurance d'indemnisation des travailleurs (code QRT : x132)**
Engagements d'assurance santé relatifs aux accidents du travail et aux maladies professionnelles, dans les cas où l'activité sous-jacente n'est pas exercée sur une base technique similaire à celle de l'assurance vie ;
4. **Assurance de responsabilité civile automobile (code QRT : x73)**
Engagements d'assurance couvrant toute responsabilité découlant de l'utilisation d'un véhicule à moteur terrestre (y compris la responsabilité du transporteur) ;
5. **Autre assurance des véhicules à moteur (code QRT : x101)**
Engagements d'assurance couvrant tout dommage subi par les véhicules à moteur terrestres (y compris les véhicules ferroviaires) ;

1. Les engagements de réassurance non-proportionnelle sont regroupés dans des segments d'activité Solvabilité 2 différents

6. **Assurance maritime, aérienne et transport (code QRT : x66)**
Engagements d'assurance couvrant tout dommage subi par les véhicules fluviaux, lacustres ou maritimes et les aéronefs, et tout dommage subi par les marchandises transportées ou les bagages quel que soit le moyen de transport. Engagements d'assurance couvrant tous les passifs découlant de l'utilisation d'aéronefs ou de navires, bateaux ou embarcations naviguant sur la mer, les lacs, rivières ou canaux (y compris la responsabilité du transporteur) ;
7. **Assurance incendie et autres dommages aux biens (code QRT : x31)**
Engagements d'assurance couvrant tout dommage subi par les biens ne relevant ni du segment d'activité x101, ni du segment d'activité x66, par suite d'un incendie, d'une explosion, des forces de la nature y compris les tempêtes, la grêle ou le gel, d'un accident nucléaire, d'un affaissement de terrain ou d'un autre événement
8. **Assurance de responsabilité civile générale (code QRT : x34)**
Engagements d'assurance couvrant tous les domaines, sauf ceux des segments d'activité x73 et x66
9. **Crédit et cautionnement (code QRT : x27)**
Engagements d'assurance couvrant l'insolvabilité, le crédit à l'exportation, la vente à tempérament, le crédit hypothécaire, le crédit agricole et les sûretés directes et indirectes ;
10. **Assurance de protection juridique (code QRT : x59)**
Engagements d'assurance couvrant les frais juridiques et le coût des actions en justice ;
11. **Assistance (code QRT : x12)**
Engagements d'assurance couvrant l'assistance aux personnes qui rencontrent des difficultés lorsqu'elles sont en voyage, en déplacement ou loin de leur lieu de résidence habituel ;
12. **Pertes pécuniaires diverses (code QRT : x70)**
Engagements d'assurance couvrant le risque d'emploi, l'insuffisance de recettes, les intempéries, la perte de bénéfices, la persistance de frais généraux, les frais commerciaux imprévus, la diminution de la valeur vénale, la perte de loyers ou de revenus, les pertes commerciales indirectes autres que celles mentionnées plus haut, les pertes pécuniaires non commerciales, ainsi que tout autre risque d'assurance non-vie qui n'est pas couvert par les segments d'activité déjà cités.

Les triangles utilisés

Pour une société et un segment d'activité donnés, les triangles renseignés dans l'état s19.01.01 sont les suivants :

- le triangle des sinistres payés bruts (non cumulés) ;
- le triangle des montants recouvrables au titre de la réassurance ;
- le triangle des sinistres payés nets (non cumulés) ;
- le triangle de la meilleure estimation des provisions pour sinistres brutes non actualisées ;
- les provisions correspondant aux montants recouvrables au titre des contrats de réassurance ;

- le triangle de la meilleure estimation des provisions pour sinistres nettes non actualisées ;
- le triangle des sinistres déclarés mais non réglés bruts (RBNS bruts) ;
- le triangle de la part de la réassurance dans les provisions déclarées dans les RBNS ;
- le triangle des sinistres déclarés mais non réglés nets de réassurance (RBNS nets).

Ces différents triangles, et plus précisément les 3 premiers et les 3 derniers de la liste précédente, ont permis de construire les triangles utilisés en entrée des différentes méthodes décrites dans le chapitre précédent.

- **le triangle des paiements bruts cumulés** construit à partir du triangle des sinistres payés bruts ;
- **le triangle des paiements nets cumulés** construit à partir du triangle des sinistres payés nets ;
- **le triangle de charges brutes** construit à partir du triangle des sinistres payés bruts et du triangle des sinistres déclarés mais non réglés bruts (RBNS bruts) ;
- **le triangle de charges nettes** construit à partir du triangle des sinistres payés nets et du triangle des sinistres déclarés mais non réglés nets (RBNS nets).

Quelques statistiques sur les données utilisées

La sélection des sociétés à considérer s'est déroulée en plusieurs étapes. Tout d'abord, toutes les sociétés présentant une provision technique non-vie non nulle ont été récupérées. Parmi ces sociétés, seules celles avec des provisions techniques non-vie supérieures à 35M€ ont été conservées. Cela représente 99% du marché français en termes de provisions techniques non-vie.

Les QRTs considérés pour ces sociétés sont les QRTs à la fin 2019. Dans le cas où le QRT à fin 2019 était absent ou non satisfaisant (i.e. les triangles disponibles étaient d'une taille inférieure à 4), le QRT à fin 2018 ou à défaut à fin 2017 a pu être récupéré. En cas de QRT non satisfaisant à fin 2017, à fin 2018 et à fin 2019, la société a finalement été exclue de l'échantillon. Du fait de leur activité particulière, les sociétés de réassurance ont également été exclues de l'échantillon. L'échantillon final représente 88% du marché français en termes de provisions techniques non-vie soit 122 milliards d'euros (G€) de réserves prudentielles (i.e. le BE de réserve ; sans le BE de primes et la marge de risque).

Le tableau suivant présente le nombre de sociétés considérées par segment d'activité, une fois les retraitements décrits précédemment effectués. A la lumière de ce tableau, trois segments d'activité ont été exclus du cadre de l'étude :

- **Indemnisation des travailleurs** : non pratiquée en France du fait du recours à la sécurité sociale ;
- **Crédit et cautionnement** : ce segment a été exclu à cause du faible nombre de données ;
- **Assurance maritime, aérienne et transport** : ce segment a été exclu car, malgré la taille raisonnable de l'échantillon initial, le marché est très concentré, trois sociétés représentant environ 85% du volume.

FIGURE 3.1 – Nombre d'organismes par segment d'activité

segment d'activité	Nombre d'organismes	BE de réserve brut
Frais médicaux	89	7,5G€
Protection du revenu	94	20G€
Indemnisation des travailleurs	3	0,3G€
Responsabilité civile automobile	70	30G€
Autre assurance des véhicules à moteur	69	3,3G€
Assurance maritime, aérienne et transport	33	1,5G€
Incendie et autres dommages aux biens	76	19G€
Responsabilité civile générale	75	38G€
Crédit et cautionnement	4	0,1G€
Protection juridique	61	1G€
Assistance	38	0,4G€
Pertes pécuniaires diverses	46	1G€

3.2 CONTRÔLE DE LA QUALITÉ DES DONNÉES

Comparaison de la provision Chain-Ladder et du BE de réserve

L'un des premiers contrôles mis en place a consisté à comparer les deux provisions suivantes :

- la provision de Chain-Ladder obtenue à partir du triangle de paiements bruts cumulés ;
- la meilleure estimation des provisions pour sinistres (*Best Estimate de sinistres*) brutes directement renseignée par l'organisme dans les états quantitatifs.

Intuitivement, ces deux provisions devraient être similaires et en tout état de cause comparables. En effet, en supposant que la méthode de Chain-Ladder soit pertinente pour le calcul des réserves, il est attendu que le montant du BE de réserves soit légèrement supérieur à la provision Chain-Ladder. Le calcul du BE de réserves tient compte de frais (frais de gestion des sinistres, frais d'administration, frais de gestion des investissements et éventuellement frais d'acquisition si non déjà payés) non pris en compte dans les triangles pour le calcul de la provision Chain-Ladder. L'actualisation réalisée pour le calcul du BE pourrait venir contrebalancer cet effet, mais reste moins prépondérante (notamment à cause des taux bas). En conséquence, sous l'hypothèse que la méthode de Chain-Ladder est pertinente pour le calcul des réserves, un écart d'environ 10% à 30% serait attendu et un écart significatif entre les deux provisions aurait pu ainsi nous permettre d'identifier facilement les point aberrants ("outliers").

Les écarts relatifs entre les deux provisions ont été étudiés pour plusieurs segments d'activité. Comme indiqué ci-dessus, il est attendu d'avoir un montant BE de sinistres supérieur à la provision Chain Ladder, soit un écart relatif positif. Le tableau 3.2 ci-dessous présente les statistiques obtenues sur la distribution des écarts relatifs, pour plusieurs segments d'activité.

Statistiques	Cas 1	Cas 2	Cas 3	Cas 4
Moyenne	140%	133%	214%	171%
1er décile	-21%	43%	25%	30%
1er quartile	-7%	78%	94%	71%
Médiane	18%	121%	152%	120%
3-ième quartile	62%	172%	276%	198%
9ième décile	141%	256%	372%	300%

FIGURE 3.2 – Écart relatif entre le BE de sinistres et la provision Chain-Ladder

Pour les quatre segments d'activité, il ressort une forte volatilité des résultats avec des fourchettes 1er décile-9ème décile assez larges :

- Pour le premier cas (segment de frais médicaux), contrairement à l'attendu (voir explication plus haut), certains écarts ressortent négatifs (1er décile et 1er quartile négatifs). En définissant un écart acceptable entre -10% et 50%, environ la moitié de l'échantillon se situe hors zone ;
- Pour les cas 2 et 3 (des segments de responsabilité civile), tous les ratios ressortent positifs sauf pour un nombre très limité de sociétés ce qui peut traduire un certain niveau de prudence mis en place pour ces segments d'activité. Nous remarquons également que les écarts peuvent parfois être très élevés pour ces segments, avec des niveaux médians à 121% et 152% (i.e. le BE ressort 2 fois plus élevé que la provision Chain-Ladder pour le niveau médian) ;
- Enfin pour le dernier cas (segment d'incendie et autres dommages aux biens), les écarts ressortent là encore très élevés, avec un niveau médian à 120%.

Ce critère, bien qu'étant intuitif, ne permet donc pas de réaliser une première sélection des données. D'autres critères ont alors été développés et sont présentés par la suite.

Critère du coefficient de détermination R^2

Pour chaque segment d'activité et chaque société, le coefficient de détermination $R^2(\hat{\lambda}_j)$ (cf. section 2.3) a été implémenté pour les trois premiers coefficients de développement ($j = 1, 2, 3$). Pour rappel, un coefficient de détermination R^2 proche de 1 traduit une stabilité du coefficient de passage entre les différentes années de souscription, et à l'inverse, un coefficient proche de 0 traduit une forte volatilité. De façon générale, un niveau faible du coefficient de détermination nécessite une analyse plus poussée du triangle considéré.

Critère de la durée

Un indicateur basé sur la durée a été également implémenté pour chaque société et chaque segment d'activité. Ce critère est basé sur les triangles de paiements (que ce soit bruts ou nets de réassurance). Pour rappel, la durée est définie comme une moyenne pondérée des dates de sortie de certains flux fi-

nanciers (dans notre cas les paiements) par ces mêmes flux :

$$\frac{\sum_{i=1}^n i * F_i}{\sum_{i=1}^n F_i}$$

où :

- i représente l'année de développement ;
- F_i représente les règlements moyens attendus pour l'année de développement i .

Dans ce choix de calcul, il est considéré que les paiements effectifs ont lieu en fin d'année. Un autre choix aurait consisté à considérer des paiements en milieu d'année (une première approximation par rapport à une hypothèse d'homogénéité des paiements tout au long de l'année). Ce choix reviendrait à pondérer les "milieux d'année" (i.e. 0.5, 1.5, etc.) par les flux F_i et le résultat aurait été diminué de 0.5. L'objectif étant de développer des indicateurs permettant d'identifier des problèmes de données, l'importance de ce choix est de second ordre.

Cette duration représente le temps moyen de liquidation d'un segment d'activité, y compris la première année. Un autre choix qui aurait pu être testé est celui de la duration des paiements des réserves. Comme une grande partie des paiements est effectuée la première année de développement (et ne rentrant donc pas dans les montants des réserves), le choix retenu est plus stable au niveau d'un segment d'activité et représente ainsi un indicateur plus fiable.

Facteur d'émergence

Le facteur d'émergence se définit comme étant le rapport entre l'écart-type obtenu par la méthode Merz & Wüthrich et l'écart-type obtenu par la méthode Mack. Pour rappel, le modèle de Mack permet d'estimer une erreur de prédiction dite "à l'ultime", là où la méthode de Merz & Wüthrich est basée sur une vision du risque à un an. Ainsi, le facteur d'émergence permet de déterminer la partie du risque qui va "émerger" à un an. Il est attendu que le facteur d'émergence soit homogène au sein d'un même segment d'activité. Les triangles avec un facteur d'émergence particulier pourrait signifier un problème dans les données ou bien signaler une société avec des risques non-homogènes avec le reste du marché.

La différence entre les écarts-types des méthodes de Merz & Wüthrich et Bootstrap

Un autre indicateur intéressant consiste à étudier la différence entre l'écart-type obtenu par la méthode de Merz & Wüthrich et l'écart-type obtenu par la méthode Bootstrap. En effet, ces deux méthodes permettent d'obtenir une mesure de la volatilité du risque de réserve à un an. De plus, les résultats sont pour la plupart du temps similaire (cf. section 3.7). Ainsi, dans le cas d'un écart

important entre ces deux écarts-types, un examen plus approfondi du triangle en entrée peut s'avérer nécessaire.

3.3 ALGORITHME DE TRAITEMENT DES DONNÉES

L'échantillon total ("triangles") de l'étude comporte 618 (cf. tableau 3.1 section 3.1). L'étude individuelle de chaque triangle aurait été très coûteuse en temps. Nous avons ainsi développé un outil qui prend en entrée les triangles de liquidation (paiements ou charges, bruts ou nets) et qui permet d'obtenir pour chaque société, chaque segment d'activité et chaque type de triangle les sorties suivantes : provision Chain-Ladder, coefficients de détermination des trois premiers coefficients de développement, durée des paiements, l'erreur de prédiction pour différentes méthodes (Mack, Merz & Wüthrich, Merz & Wüthrich simplifié, bootstrap).



FIGURE 3.3 – Entrées et sorties de l'outil développé

Le traitement des sorties s'est déroulé en plusieurs étapes :

1. Regroupement des résultats par segment d'activité ;
2. Identification des outliers à partir des indicateurs développés ;
3. Analyse individuelle de chaque outlier (coefficients de développement, résidus, sigma de Mack, etc.) ;
4. Après analyse, plusieurs choix ont pu être faits :
 - Si après examen, aucune "anomalie" particulière n'a été identifiée, la société est conservée dans l'échantillon. Par "anomalie", nous comprenons une des raisons de retraitement et qui sont toutes expliquées dans la section 3.4 suivante ;
 - Si une "anomalie" a été identifiée, alors le triangle a été retraité (les détails sont également donnés dans la section suivante). Il est à noter qu'aucun montant n'a été modifié, seules des années de survenance ou des années civiles (diagonales) ont été supprimées ;
 - Si plusieurs "anomalies" ont été identifiées, le "triangle" a été gardé dans le cadre de l'étude tant que les retraitements ne mènent pas à une diminution trop importante de sa taille (longueur) ;
5. Afin de définir l'échantillon final utilisé pour l'étude, des "filtres" ont été utilisés et qui seront indiqués (et justifiés) pour chaque segment d'activité (volume trop faible, profondeur d'historique insuffisante, durée atypique).

Le retraitement des données s'avère délicat. En effet, le but n'est pas d'exclure un triangle "atypique" juste parce que son coefficient de développement (quelle que soit la mesure utilisée) ne se comporte pas de la manière attendue. Il est tout à fait possible que les risques représentés dans un "triangle" ne soient pas homogènes. L'information que nous avons à disposition est au niveau du segment de risque alors qu'il peut être composé de plusieurs groupes de risques homogènes. Comme nous n'avons pas cette information, exclure ou non des triangles car les résultats sont considérés "atypiques" deviendrait ainsi aléatoire. Nous préférons ainsi n'exclure qu'une partie du triangle dont la raison pourrait être identifiée (devinée) de l'extérieur et sans connaissance particulière de la société.

3.4 RETRAITEMENT DES DONNÉES

Comme indiqué dans la précédente section, quand certains indicateurs statistiques sont atypiques, nous avons analysé les données (i.e. les triangles) individuellement. Nous avons identifié plusieurs raisons qui pourraient perturber les triangles de manière à les rendre inutilisables :

- rupture dans le provisionnement ;
- année atypique ;
- acquisition ou vente de portefeuille ;
- forte croissance ou baisse du chiffre d'affaires ;
- erreurs manifestes de données.

Rupture dans le provisionnement

Une des raisons qui peut perturber les "triangles" de manière à ce que la volatilité ne reflète plus le vrai profil de risque de la société est une rupture dans le provisionnement. Cela peut résulter :

- de changements dans le processus de règlement des sinistres ou du traitement des sinistres ;
- un changement de législation qui a un impact sur les coûts des sinistres individuels ;

La rupture dans le provisionnement se traduit quantitativement par un changement brusque, d'une année sur l'autre, des coefficients de développement individuel. Ci-dessous est donné en exemple un triangle de paiements (figure 3.4) sur une branche courte où seul le premier coefficient de développement a un impact sur le niveau des provisions, ainsi que le triangle des coefficients de développement individuel (figure 3.5).

Nous observons que les trois premières années de survenance sont caractérisées par des coefficients de développement individuels d'environ 1,12 alors qu'ensuite les coefficients de développement individuels sont compris entre 1,27 et 1,31 (cf. figure 3.5). Le tableau ci-dessous 3.6 indique, en fonction de la profondeur d'historique, le coefficient de variation de Merz & Wüthrich (variante simplifiée) ainsi que le coefficient de détermination du premier coefficient de développement individuel $R2(\hat{\lambda}_1)$.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	926	1042	1045	1045	1045	1045	1045	1045	1045	1045	1045	1045	1045	1045
2	936	1044	1046	1046	1046	1046	1046	1046	1046	1046	1046	1046	1046	
3	967	1091	1094	1094	1094	1094	1094	1094	1094	1094	1094	1094	1094	
4	878	1134	1137	1137	1137	1137	1137	1137	1137	1137	1137	1137		
5	927	1175	1178	1178	1178	1178	1178	1178	1178	1178				
6	906	1165	1167	1167	1167	1167	1167	1167	1167					
7	961	1229	1233	1233	1233	1233	1233	1233						
8	972	1252	1257	1257	1257	1257	1257							
9	1081	1382	1387	1387	1387	1387								
10	1096	1404	1408	1408	1408									
11	1093	1393	1396	1397										
12	1045	1363	1367											
13	1024	1326												
14	1000													

FIGURE 3.4 – Exemple de rupture dans le provisionnement - triangle de paiements

	1	2	3	4	5	6	7	8	9	10	11	12	13
1	1,12	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
2	1,11	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	
3	1,13	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00		
4	1,29	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00			
5	1,27	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00				
6	1,29	1,00	1,00	1,00	1,00	1,00	1,00	1,00					
7	1,28	1,00	1,00	1,00	1,00	1,00	1,00						
8	1,29	1,00	1,00	1,00	1,00	1,00							
9	1,28	1,00	1,00	1,00	1,00								
10	1,28	1,00	1,00	1,00									
11	1,27	1,00	1,00										
12	1,30	1,00											
13	1,30												

FIGURE 3.5 – Exemple de rupture dans le provisionnement - les coefficients de développement individuel

Historique	14	13	12	11	10	9	8	7	6
σ_{MW} simplifié	28,4%	24,7%	17,8%	3,91%	4,07%	3,73%	4,04%	4,30%	4,85%
$R2(\hat{\lambda}_1)$	76,9%	77,8%	83,5%	98,9%	98,6%	98,6%	97,5%	96,8%	94,5%

FIGURE 3.6 – Exemple de rupture dans le provisionnement - valeurs de σ_{MW} et de $R2(\hat{\lambda}_1)$ en fonction de la profondeur d'historique

L'évolution des coefficients de variation σ_{MW} en fonction de la taille du "triangle" utilisé indique bien une rupture dans le provisionnement qui se traduit par une augmentation significative (et probablement induite) de la volatilité. Ce changement est également visible sur le coefficient de détermination $R2(\hat{\lambda}_1)$. Dans des cas comme celui-ci, afin de rajouter un maximum de données à notre étude, nous utilisons un "triangle" de plus petite taille afin d'éliminer la rupture dans le provisionnement. Plus précisément, dans l'exemple *supra*, nous supprimons les trois premières années de survenance. Dans le cas où la rupture de provisionnement est plus récente et la taille du triangle retraité devient trop petite, nous faisons le choix de ne pas garder le "triangle" pour notre étude.

Remarque 3.4.1 Cet exemple n'est pas suffisant pour généraliser sur le niveau adéquat du coefficient de détermination $R2$. En effet, la valeur critique

du R^2 n'est pas évidente à fixer et, en tout état de cause, elle est fortement liée au segment d'activité. Dans cet exemple, les données proviennent du segment *frais médicaux* dont le paiement des sinistres est effectué pendant les deux premières années de développement (branche courte). Le niveau des provisions est donc lié au premier coefficient de passage, qui traduit le pourcentage des paiements effectués lors de la deuxième année par rapport à la première année (de survenance des sinistres). Un coefficient de détermination R^2 faible se traduit donc mécaniquement par un fort effet sur la volatilité. Intuitivement, ceci n'est pas toujours le cas pour des segments à développement long comme ceux de responsabilité civile car la volatilité n'est pas expliquée par un seul coefficient de passage (ou dit autrement, une seule année de développement ne suffit pas pour connaître la charge à l'ultime et la volatilité du provisionnement).

Année atypique

Un autre cas qui peut perturber un "triangle" de manière à ce que la volatilité ne reflète pas le vrai profil de risque de la société est celui d'une année atypique. C'est par exemple le cas de l'année 2020 : la crise sanitaire COVID-19 en fait une année atypique. En effet, à cause du confinement et des problèmes opérationnels que les sociétés d'assurance ont connus, les paiements des triangles ont pu être ralentis, ce qui se traduira dans les prochains exercices par des paiements plus importants et donc des coefficients de passage individuel plus significatifs que d'habitude. Cependant, cette particularité ne traduit pas forcément une volatilité de la charge à l'ultime. Un autre cas typique a été celui des tempêtes hivernales (Lothar et Martin) de la fin de l'année 1999. La majorité des sinistres étant déclarée en 2000, le premier coefficient de passage pour l'année de survenance 1999 a été particulièrement élevé par rapport à la moyenne. A nouveau, cet événement ne traduit pas forcément une volatilité de la charge à l'ultime.

Sans prendre en compte des événements exceptionnels comme les exemples ci-dessus, il est tout à fait possible d'avoir d'autres raisons pour avoir des coefficients de développement individuel atypiques. Cela peut être tout simplement provenir d'un sinistre très coûteux, des problèmes opérationnels ou d'une toute autre raison qui accélère ou ralentit le paiement des sinistres. Dans une société d'assurance, les actuaires sont en mesure de contacter les gestionnaires sinistres pour comprendre l'origine de ces effets, ce qui n'a pas été possible lors de notre étude. Nous essayons donc de déduire les éventuelles années atypiques grâce à des indicateurs statistiques comme le coefficient de détermination R^2 .

Dans le triangle ci-dessous (figure 3.7), le premier coefficient de développement de la quatrième année de survenance (i.e. $\lambda_{4,1} = 2,04$ cf. 3.8) se distingue, alors que la moyenne des autres coefficients individuels de la première colonne est autour de 2,65. Il est d'ailleurs facile d'observer que pour la première année de développement de cette année de survenance, les règlements ont été plus importants que d'habitude (probablement du fait d'un ou plusieurs gros sinistres). Dans ce cas, nous retirons la quatrième année de survenance car elle

est très probablement atypique, et réduisons ainsi la dimension du triangle.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	892	2 392	2 960	3 312	3 689	3 832	4 347	4 410	4 458	4 503	4 531	4 541	4 532	4 551
2	925	2 436	3 052	3 340	3 739	3 813	4 012	4 105	4 158	4 454	4 516	4 612	4 859	
3	962	2 406	2 978	3 324	3 541	3 742	3 884	3 994	4 391	4 496	4 563	5 376		
4	1 514	3 090	3 844	4 180	4 470	4 895	5 017	5 125	5 156	5 277	5 529			
5	963	2 564	3 107	3 630	3 817	4 004	4 093	4 374	5 217	5 903				
6	1 046	2 825	3 643	3 912	4 174	4 403	4 883	4 996	5 256					
7	1 113	2 776	3 474	3 887	4 876	5 102	5 247	5 303						
8	935	2 523	3 321	3 900	4 663	4 813	5 082							
9	984	2 619	3 323	3 631	4 247	4 364								
10	823	2 276	2 890	3 192	3 543									
11	798	2 109	2 754	3 188										
12	912	2 375	3 092											
13	885	2 381												
14	1 000													

FIGURE 3.7 – Exemple d'année atypique - triangle de paiements

	1	2	3	4	5	6	7	8	9	10	11	12	13
1	2,68	1,24	1,12	1,11	1,04	1,13	1,01	1,01	1,01	1,01	1,00	1,00	1,00
2	2,63	1,25	1,09	1,12	1,02	1,05	1,02	1,01	1,07	1,01	1,02	1,05	
3	2,50	1,24	1,12	1,07	1,06	1,04	1,03	1,10	1,02	1,01	1,18		
4	2,04	1,24	1,09	1,07	1,09	1,02	1,02	1,01	1,02	1,05			
5	2,66	1,21	1,17	1,05	1,05	1,02	1,07	1,19	1,13				
6	2,70	1,29	1,07	1,07	1,05	1,11	1,02	1,05					
7	2,49	1,25	1,12	1,25	1,05	1,03	1,01						
8	2,70	1,32	1,17	1,20	1,03	1,06							
9	2,66	1,27	1,09	1,17	1,03	-							
10	2,76	1,27	1,10	1,11									
11	2,64	1,31	1,16										
12	2,60	1,30											
13	2,69												

FIGURE 3.8 – Exemple d'année atypique - les coefficients de développement individuel

À titre de comparaison, avant retraitement $R2(\hat{\lambda}_1) = 2,9\%$ et la valeur du coefficient de variation de Merz & Wüthrich (variante simplifiée) $\sigma_{MW} = 15,31\%$ alors qu'après retraitement $R2(\hat{\lambda}_1) = 84,2\%$ et la valeur du coefficient de variation de Merz & Wüthrich (variante simplifiée) $\sigma_{MW} = 9,69\%$. Nous observons ainsi une amélioration significative de l'indicateur $R2$ qui se traduit également par une baisse du coefficient de variation.

Acquisition ou vente de portefeuille

L'acquisition et la vente d'un portefeuille d'assurance sont aussi de nature à changer fortement les coefficients de développement individuel d'une année "civile" (i.e. diagonale) et in fine la volatilité d'un "triangle". Comme indiqué dans la section précédente, nous n'avons pas cette information, mais la nature du triangle suivant (figure 3.9) peut fortement laisser à penser à ce cas de figure.

	1	2	3	4	5	6	7	8	9	10	11	12
1	290	502	578	579	610	623	624	630	648	648	645	646
2	399	676	724	735	759	761	765	771	773	763	764	
3	573	981	1 078	1 104	1 106	1 113	1 121	1 124	1 119	1 116		
4	634	1 212	1 374	1 410	1 418	1 441	1 464	1 468	1 460			
5	488	934	1 038	1 052	1 077	1 081	1 086	1 083				
6	278	608	655	683	689	1 091	1 084					
7	397	864	1 081	1 093	2 118	2 117						
8	364	835	917	2 180	2 412							
9	364	877	2 205	2 462								
10	469	1 997	2 976									
11	725	4 084										
12	1 000											

FIGURE 3.9 – Exemple d'acquisition de portefeuille - triangle de paiements

En regardant par année de développement, les coefficients de développement individuel entre les années “civiles” $n - 2$ et $n - 1$ (la dernière diagonale étant l'année n) ressortent complètement différents des autres (cf. tableau 3.10) :

- sur la première colonne, le coefficient $\hat{\lambda}_{10,1} = 4,26$ alors que la moyenne des coefficients de développement individuel $\hat{\lambda}_{i,1}, i = 1, \dots, 9$ est égale à 2 ;
- sur la deuxième colonne, le coefficient $\hat{\lambda}_{9,2} = 2,51$ alors que la moyenne des coefficients de développement individuel $\hat{\lambda}_{i,2}, i = 1, \dots, 8$ est égale à 1,12 ;
- sur la troisième colonne, le coefficient $\hat{\lambda}_{8,3} = 2,38$ alors que la moyenne des coefficients de développement individuel $\hat{\lambda}_{i,3}, i = 1, \dots, 7$ est égale à 1,02 ;
- sur la quatrième colonne, le coefficient $\hat{\lambda}_{7,4} = 1,94$ alors que la moyenne des coefficients de développement individuel $\hat{\lambda}_{i,4}, i = 1, \dots, 6$ est égale à 1,02 ; sur la cinquième colonne, le coefficient $\hat{\lambda}_{6,5} = 1,58$ alors que la moyenne des coefficients de développement individuel $\hat{\lambda}_{i,5}, i = 1, \dots, 5$ est égale à 1,01.

	1	2	3	4	5	6	7	8	9	10	11
1	1,73	1,15	1,00	1,05	1,02	1,00	1,01	1,03	1,00	0,99	1,00
2	1,69	1,07	1,01	1,03	1,00	1,00	1,01	1,00	0,99	1,00	
3	1,71	1,10	1,02	1,00	1,01	1,01	1,00	1,00	1,00		
4	1,91	1,13	1,03	1,01	1,02	1,02	1,00	0,99			
5	1,92	1,11	1,01	1,02	1,00	1,00	1,00				
6	2,19	1,08	1,04	1,01	1,58	0,99					
7	2,17	1,25	1,01	1,94	1,00						
8	2,29	1,10	2,38	1,11							
9	2,41	2,51	1,12								
10	4,26	1,49									
11	5,64										

FIGURE 3.10 – Exemple d'acquisition de portefeuille - les coefficients de développement individuel

Dans ce cas bien précis, nous supprimons les deux dernières diagonales. Si ce phénomène était apparu sur une diagonale inférieure, par exemple la troisième, nous aurions effacé les trois premières années de survenance. A titre de comparaison, avant retraitement les coefficients de détermination des trois pre-

miers coefficients de développement étaient de $R2(\hat{\lambda}_1) = 55,9\%$, $R2(\hat{\lambda}_2) = 74\%$ et respectivement $R2(\hat{\lambda}_3) = 61,0\%$ alors que la valeur du coefficient de variation de Merz & Wüthrich (variante simplifié) $\sigma_{MW} = 36,92\%$. Après retraitement, nous avons $R2(\hat{\lambda}_1) = 80,3\%$, $R2(\hat{\lambda}_2) = 96,9\%$ et respectivement $R2(\hat{\lambda}_3) = 99,9\%$ alors que la valeur du coefficient de variation de Merz & Wüthrich (variante simplifié) $\sigma_{MW} = 17,4\%$. Nous observons ainsi une amélioration significative des indicateurs $R2$ qui se traduit également par une baisse du coefficient de variation.

Un effet similaire est observable en cas de vente de portefeuille, avec la seule différence que dans ce cas les coefficients de développement individuel sont inférieurs à 1. Cet effet est expliqué par le fait qu'à la suite de la vente du portefeuille, les paiements relatifs aux sinistres affiliés à des contrats sortants ne sont plus "affichés" dans le triangle. Ainsi, dès l'année suivante, la société peut analyser l'évolution des seuls sinistres en portefeuille. Dans l'exemple suivant (figure 3.11), nous observons une baisse des règlements cumulés entre les diagonales $n - 4$ et $n - 3$ (sauf le tout premier point de la diagonale).

	1	2	3	4	5	6	7	8	9	10	11	12	13
1	806	2 478	3 023	3 198	3 382	3 547	3 599	3 721	3 757	3 769	3 767	3 769	3 773
2	963	2 737	3 171	3 475	3 833	4 035	4 189	4 425	4 531	3 383	3 395	3 533	
3	1 058	2 702	3 160	3 439	3 628	3 751	3 836	4 119	2 992	3 008	3 049		
4	1 321	3 430	4 043	4 392	4 630	4 803	4 888	3 942	3 946	4 020			
5	1 387	3 406	4 192	4 505	4 811	4 896	4 183	4 293	4 285				
6	1 223	3 214	3 786	4 125	4 271	3 510	3 701	3 757					
7	1 322	3 300	4 085	4 438	3 635	3 835	4 036						
8	969	3 515	4 006	3 495	3 759	3 638							
9	1 530	3 918	3 633	3 771	3 943								
10	1 469	3 076	3 357	3 641									
11	1 233	3 428	3 904										
12	1 151	3 427											
13	1 000												

FIGURE 3.11 – Exemple de vente de portefeuille - triangle de paiements

Cet effet est plus facile à remarquer à partir du triangle des coefficients de développement individuel ci-dessous (figure 3.12) :

- sur la première colonne, le coefficient $\hat{\lambda}_{10,1} = 2,09$ est bien plus petit que la moyenne des coefficients individuels $\hat{\lambda}_{i,1}, i = 1, \dots, 9$ est égale à 2,76 ;
- sur la deuxième colonne, le coefficient $\hat{\lambda}_{9,2} = 0,93 < 1$ alors que la moyenne des coefficients individuels $\hat{\lambda}_{i,2}, i = 1, \dots, 8$ est égale à 1,19 ;
- sur la troisième colonne, le coefficient $\hat{\lambda}_{8,3} = 0,87 < 1$ alors que la moyenne des coefficients individuels $\hat{\lambda}_{i,3}, i = 1, \dots, 7$ est égale à 1,08 ;
- le même type d'effet est observable sur l'ensemble de la diagonale pour les coefficients individuels $\hat{\lambda}_{7,4}, \hat{\lambda}_{6,5}, \dots, \hat{\lambda}_{2,9}$.

Dans ce cas de figure, nous effaçons les trois dernières diagonales. A titre de comparaison, avant retraitement les coefficients de détermination des trois premiers coefficients de développement étaient de $R2(\hat{\lambda}_1) = 46,2\%$, $R2(\hat{\lambda}_2) = 58,3\%$ et respectivement $R2(\hat{\lambda}_3) = 70,8\%$ alors que la valeur du coefficient de variation de Merz & Wüthrich (variante simplifié) $\sigma_{MW} = 36,92\%$. Après retraitement, les valeurs du $R2$ sont $R2(\hat{\lambda}_1) = 64,5\%$, $R2(\hat{\lambda}_2) = 94,3\%$ et respectivement $R2(\hat{\lambda}_3) = 100\%$ alors que la valeur du coefficient de variation de Merz & Wü-

	1	2	3	4	5	6	7	8	9	10	11	12
1	3,07	1,22	1,06	1,06	1,05	1,01	1,03	1,01	1,00	1,00	1,00	1,00
2	2,84	1,16	1,10	1,10	1,05	1,04	1,06	1,02	0,75	1,00	1,04	
3	2,55	1,17	1,09	1,05	1,03	1,02	1,07	0,73	1,01	1,01		
4	2,60	1,18	1,09	1,05	1,04	1,02	0,81	1,00	1,02			
5	2,46	1,23	1,07	1,07	1,02	0,85	1,03	1,00				
6	2,63	1,18	1,09	1,04	0,82	1,05	1,02					
7	2,50	1,24	1,09	0,82	1,05	1,05						
8	3,63	1,14	0,87	1,08	0,97							
9	2,56	0,93	1,04	1,05								
10	2,09	1,09	1,08									
11	2,78	1,14										
12	2,98											

FIGURE 3.12 – Exemple de vente de portefeuille - les coefficients de développement individuel

thrich (variante simplifié) $\sigma_{MW} = 8,75\%$. Nous observons ainsi une amélioration significative des indicateurs R^2 qui se traduit également par une baisse très importante du coefficient de variation.

Remarque 3.4.2 Les deux derniers exemples sont également intéressants d'un autre point de vue, le niveau des coefficients de variation ne sont pas parfaitement corrélés aux valeurs des coefficients de détermination R^2 (i.e. à la qualité de la régression). Pour le cas de l'acquisition de portefeuille (triangles 3.9 et 3.10), la qualité de la régression des trois premiers facteurs de développement après retraitement est bien supérieure à ceux du cas de la vente de portefeuille (triangles 3.11 et 3.12). En revanche, le coefficient de variation du premier cas (i.e. acquisition) reste bien supérieur à celui du deuxième cas (i.e. vente). Ceci peut s'expliquer par le fait que le risque sous-jacent n'est pas le même (plus précisément deux segments d'activité différents) mais aussi par les volumes réels de chaque société (nous rappelons qu'en utilisant le lemme de la section 2.7 nous anonymisons les données).

Forte croissance ou baisse du chiffre d'affaires

Une forte croissance (cf. figure 3.13) ou une forte baisse du chiffre d'affaires (cf. figure 3.13) donne lieu la plupart du temps à une volatilité importante. Les raisons peuvent être multiples, dont certaines sont rappelées ci-dessous :

- quand les volumes sont faibles, la loi des grands nombres n'est pas vérifiée ce qui donne lieu à des coefficients de développement individuel plus volatiles ;
- la volatilité de ces coefficients de développement individuel peut également être expliquée par des raisons opérationnelles (évolution des services de gestion, de la politique de souscription, de la politique de provisionnement etc.) ;
- changement/évolution de la population assurée se traduisant par des risques hétérogènes.

Les deux cas de croissance et respectivement baisse forte du chiffre d'affaires *infra* sont des exemples dans ce sens. À titre d'information, les coefficients

de variation de Merz & Wüthrich (variante simplifiée) sont respectivement de 42,03% et de 99,27%.

	1	2	3	4	5	6	7	8	9	10
1	2	2	2	2	2	2	2	2	2	2
2	2	2	2	2	2	2	2	2	2	
3	3	4	4	4	4	4	4	4		
4	7	12	13	13	13	13	13			
5	15	27	28	28	28	28				
6	69	95	97	97	97					
7	66	159	164	167						
8	352	443	447							
9	463	539								
10	1 000									

FIGURE 3.13 – Exemple de croissance forte - triangle de paiements

	1	2	3	4	5	6	7	8	9	10	11	12
1	17 214	22 016	22 094	22 109	22 112	22 112	22 112	22 112	22 112	22 112	22 112	22 112
2	19 359	24 026	24 522	24 541	24 635	24 643	24 690	24 696	24 426	24 440	24 722	
3	18 079	21 328	21 394	21 466	21 481	21 486	21 486	21 486	21 486	21 486		
4	10 649	14 345	14 511	14 679	14 796	15 012	15 064	15 076	15 079			
5	6 814	9 466	9 526	9 521	9 524	9 521	9 528	9 528				
6	4 121	5 704	5 751	5 777	5 784	5 792	5 793					
7	1 699	3 098	3 277	3 317	3 319	3 323						
8	841	1 535	1 814	1 828	1 844							
9	864	2 235	2 417	2 467								
10	895	1 816	2 099									
11	1 256	2 886										
12	1 000											

FIGURE 3.14 – Exemple de décroissance forte - triangle de paiements

Quand nous identifions ce type de triangle, nous les enlevons de l'échantillon de notre étude même si les indicateurs statistiques $R2$ ou les coefficients de variation paraissent raisonnables.

Remarque 3.4.3 Il est cependant à noter que les indicateurs statistiques pour la qualité des coefficients de développement ne sont pas fiables dans ces cas. En effet, comme ils sont pondérés par les paiements (dans ces exemples, et de manière générale par les montants $C_{i,j}$), un poids important est donné à peu d'années de survenance ce qui se traduit par des indicateurs $R2$ de "bonne qualité" même si les coefficients de variation sont élevés. Dans les exemples ci-dessus, nous avons :

- pour l'exemple avec la forte croissance du chiffre d'affaires 3.13, $R2(\hat{\lambda}_1) = 98,3\%$ et le coefficient de variation de Merz & Wüthrich (variante simplifiée) $\sigma_{MW} = 42,03\%$;
- pour l'exemple avec la forte baisse du chiffre d'affaires 3.14, $R2(\hat{\lambda}_1) = 99,7\%$ et le coefficient de variation de Merz & Wüthrich (variante simplifiée) $\sigma_{MW} = 99,27\%$.

Remarque 3.4.4 Les deux exemples ci-dessus permettent de comprendre que les indicateurs statistiques $R2$ ne suffisent pas à identifier les problèmes

de qualité de données. C'est la raison principale qui a motivé le développement d'autres indicateurs. Alors que la méthode *bootstrap* avec une loi normale pour simuler la nouvelle diagonale donne globalement des résultats proches du modèle de Merz & Wüthrich (cette affirmation sera montrée lors de la section suivante 3.6), dans le cas de l'exemple sur la forte croissance du chiffre d'affaires, l'écart des coefficients de variation est d'environ 35%. En ce qui concerne le cas de l'exemple sur la forte baisse du chiffre d'affaires, l'écart des coefficients de variation est d'environ 10% (dans l'autre sens).

Erreurs de données

Un problème souvent rencontré est l'erreur dans les données dont un exemple pratique est donné par le "triangle" ci-dessous (figure 3.15). Même si l'évolution du chiffre d'affaires est significative (sans connaître le montant des primes, il semblerait que le chiffre d'affaires double en 12 ans), sa croissance est globalement stable. Les montants (paiements cumulés bruts de réassurance) sont du même ordre de grandeur (globalement entre 500 et 1500) à l'exception de $C_{1,1} = 35$. Nous sommes de toute évidence dans le cas d'une erreur ou d'une perte sur les données.

	1	2	3	4	5	6	7	8	9	10	11	12
1	35	518	556	571	584	608	622	648	675	698	707	716
2	493	791	825	830	835	842	850	853	853	853		
3	767	1090	1172	1190	1210	1243	1250	1266	1269	1270		
4	686	959	985	989	993	994	998	998	1003			
5	548	910	955	1022	1075	1134	1180	1216				
6	615	904	947	965	980	1006	1022					
7	612	1051	1117	1129	1134	1135						
8	700	1181	1243	1256	1266							
9	628	969	1016	1032								
10	914	1373	1453									
11	963	1476										
12	1000											

FIGURE 3.15 – Exemple d'erreur dans les données - triangle de paiements

Cela s'observe également sur le triangle des coefficients de développement individuel (figure 3.16). Alors que les coefficients de développement individuel sont entre 1 et 2, le tout premier est bien plus grand $\hat{\lambda}_{1,1} = 14,77$. Cela se traduit par un coefficient de variation de Merz & Wüthrich (variante simplifiée) $\sigma_{MW} = 135\%$.

Dans l'exemple ci-dessus, nous supprimons l'année de survenance qui concerne l'erreur. Dans le cas où plusieurs années de survenance sont concernées par des erreurs de ce type, le triangle en question est exclu de l'échantillon de notre étude.

Remarque 3.4.5 Comme pour les deux exemples précédents sur la forte croissance et décroissance du chiffre d'affaires, les indicateurs statistiques $R2$ sont globalement de bonne qualité ($R2(\hat{\lambda}_1) = 88,7\%$, $R2(\hat{\lambda}_2) = 99,7\%$ et $R2(\hat{\lambda}_3) = 99,1\%$) alors que le coefficient de variation est élevé (pour rappel $\sigma_{MW} = 135\%$). En revanche, comme dans les deux exemples précédents, l'écart

	1	2	3	4	5	6	7	8	9	10	11
1	14,77	1,07	1,03	1,02	1,04	1,02	1,04	1,04	1,03	1,01	1,01
2	1,61	1,04	1,01	1,01	1,01	1,01	1,00	1,00	1,00	1,00	
3	1,42	1,08	1,02	1,02	1,03	1,01	1,01	1,00	1,00		
4	1,40	1,03	1,00	1,00	1,00	1,00	1,00	1,00			
5	1,66	1,05	1,07	1,05	1,06	1,04	1,03				
6	1,47	1,05	1,02	1,02	1,03	1,02					
7	1,72	1,06	1,01	1,00	1,00						
8	1,69	1,05	1,01	1,01							
9	1,54	1,05	1,02								
10	1,50	1,06									
11	1,53										

FIGURE 3.16 – Exemple d'erreur dans les données - les coefficients de développement individuel

avec le coefficient de variation estimé avec la méthode de Bootstrap est un bon indicateur des problèmes sur les données qui, dans cet exemple, est environ 20 fois plus grand.

Exemple de mauvaise qualité de données

Dans certains cas les données sont de mauvaise qualité, statistique en tout état de cause, qui rend impossible leur utilisation. C'est notamment le cas de l'exemple suivant (figure 3.17) où les montants évoluent d'une manière erratique.

	1	2	3	4	5	6	7	8	9	10	11	12
1	690	2 470	2 727	2 791	2 797	2 797	2 812	2 814	1 859	1 859	1 859	1 862
2	276	553	575	578	588	591	594	595	607	609	609	
3	888	1 776	2 840	3 091	3 261	3 264	3 306	3 310	3 099	3 099		
4	406	742	778	784	774	771	772	764	765			
5	538	868	915	923	929	935	935	938				
6	678	1 203	1 014	1 066	1 073	926	928					
7	452	855	972	984	996	997						
8	1 503	1 925	1 995	2 028	2 044							
9	475	1 279	1 372	1 395								
10	435	882	940									
11	428	857										
12	1 000											

FIGURE 3.17 – Exemple de mauvaise qualité de données - triangle de paiements

Cela s'observe de manière plus flagrante sur le triangle de coefficients de développement individuel (figure 3.18) :

- sur la première colonne, la moyenne des coefficients individuels est autour de 2 alors que trois coefficients sont très différents $\hat{\lambda}_{1,1} = 3,58$, $\hat{\lambda}_{8,1} = 1,29$ et $\hat{\lambda}_{9,1} = 2,69$;
- sur la deuxième colonne, la moyenne des coefficients individuels est légèrement supérieure à 1 alors que deux coefficients sont très différents $\hat{\lambda}_{3,2} = 1,60$ et $\hat{\lambda}_{3,2} = 0,84$;
- sur la quatrième colonne, la moyenne des coefficients individuels est autour de 1 alors qu'un coefficient vaut $\hat{\lambda}_{3,2} = 0,86$;

- sur la cinquième colonne, la moyenne des coefficients individuels est autour de 1 alors qu'un coefficient vaut $\hat{\lambda}_{3,2} = 0,66$.

	1	2	3	4	5	6	7	8	9	10	11
1	3,58	1,10	1,02	1,00	1,00	1,01	1,00	0,66	1,00	1,00	1,00
2	2,00	1,04	1,01	1,02	1,00	1,01	1,00	1,02	1,00	1,00	
3	2,00	1,60	1,09	1,05	1,00	1,01	1,00	0,94	1,00		
4	1,83	1,05	1,01	0,99	1,00	1,00	0,99	1,00			
5	1,61	1,05	1,01	1,01	1,01	1,00	1,00				
6	1,77	0,84	1,05	1,01	0,86	1,00					
7	1,89	1,14	1,01	1,01	1,00						
8	1,28	1,04	1,02	1,01							
9	2,69	1,07	1,02								
10	2,03	1,07									
11	2,00										

FIGURE 3.18 – Exemple de mauvaise qualité de données - les coefficients de développement individuel

Dans ce cas de figure, le retraitement de données a peu de sens : il est ainsi préférable d'exclure le triangle de l'échantillon de l'étude.

3.5 TRIANGLES DE CHARGES INEXPLOITABLES

Avant de passer au chapitre suivant montrant les résultats pratiques de notre étude, nous avons malheureusement constaté que les triangles de charges ne sont pas exploitables. Ce constat a été effectué sur l'ensemble des segments d'activité et les raisons sont multiples :

1. **Triangles "RBNS" vides** : Un nombre significatif d'entreprises ne renseignent pas les triangles de "RBNS" ce qui en pratique revient à avoir des triangles de charges égaux aux triangles de paiements. Le nombre de triangles "RBNS" manquants n'est pas homogène entre les segments d'activité², mais à titre d'exemple pour le segment d'activité de frais médicaux, ce problème est rencontré sur 25% de l'échantillon (plus précisément, 25 sociétés sur un total de 89) ;
2. **La prudence du provisionnement dossier/dossier** : Le provisionnement du dossier-dossier est réputé être prudent sur le marché français notamment en cas de sinistres pouvant coûter chers et être litigieux. Cela se traduit en pratique par des baisses régulières et significatives sur les triangles "RBNS", qui ne sont pas expliquées par les paiements, de nature à perturber l'homogénéité des triangles de charges. A titre d'exemple, le tableau suivant montre les écarts relatifs du risque à l'ultime (modèle de Mack) sur le segment "autre assurance des véhicules à moteur" entre les triangles de paiements bruts et les triangles de charges brutes.

Nous constatons que globalement les triangles de charges conduisent à une bien plus forte volatilité (+80% en moyenne) alors que le segment d'activité est relativement court. Le coût d'un sinistre reste cependant limité en fonction du prix du véhicule (très souvent une voiture) et il est globalement facile à anticiper.

2. pour les segments de responsabilité civile l'ensemble des triangles "RBNS" sont disponibles

Statistique	Écart relatif
Moyenne	80%
1er décile	-27%
1er quartile	5%
3-ième quartile décile	117%
9ième décile	216%

FIGURE 3.19 – Écart relatif entre le risque à l'ultime entre les triangles de charges et les triangles de paiements bruts pour le segment "Autre assurance des véhicules à moteur"

3. **Règles de provisionnement évolutives** : Les règles et les orientations du provisionnement dossier-dossier, qui sont à la base des triangles "RBNS", peuvent changer au sein de l'entreprise au cours des années, ce qui rend les triangles plus volatiles. Notamment, le niveau de prudence (cf. point précédent) peut évoluer. Cela n'est pas le cas des triangles de paiements car ils ne représentent pas une anticipation des gestionnaires sinistres mais bien des paiements réels.
4. **Écart entre le niveau des réserves** : Le niveau des réserves estimé à partir du triangle des charges ou du triangle des paiements devrait être relativement faible notamment pour les segments à liquidation rapide. Le tableau ci-dessous montre cependant des écarts significatifs pour le même segment "autre assurance des véhicules à moteur" :

Statistique	Écart relatif
Moyenne	16%
1er décile	-57%
1er quartile	-20%
3-ième quartile décile	33%
9ième décile	89%

FIGURE 3.20 – Écart relatif des réserves en fonction du triangle utilisé (charges brutes vs. paiements bruts) pour le segment "Autre assurance des véhicules à moteur"

Nous constatons que la plupart du temps les triangles de charges conduisent à un niveau des réserves plus important (+16% en moyenne). Même si l'écart moyen reste relativement faible pour environ un tiers de l'échantillon, l'écart va du simple au double (dans un sens ou d'un autre), ce qui a une influence significative sur le calcul du coefficient de variation.

5. **Cohérence avec le calibrage de la formule standard** : Lors du calibrage de la formule standard, EIOPA a utilisé également des triangles de paiements bruts. Comme nous voulons comparer nos résultats avec ceux de la formule standard, nous privilégions dans cette étude les triangles de paiements.

3.6 DIFFÉRENCE ENTRE LES ÉCARTS-TYPES DES MÉTHODES DE MERZ & WÜTHRICH ET BOOTSTRAP

Un des indicateurs étudiés est la différence entre l'écart-type obtenu par la méthode de Merz & Wüthrich et l'écart-type obtenu par la méthode de Boots-

trap. Pour rappel, nous avons vu que dans certains cas, par exemple dans le cas d'une forte croissance ou d'une chute du chiffre d'affaires, les indicateurs statistiques R^2 ne permettaient pas d'identifier les problèmes dans les données. A l'inverse, dans ces situations, l'écart entre les coefficients de variation Merz & Wüthrich et Bootstrap (avec simulation par loi normale) ressort important, ce qui en fait un bon indicateur.

La présente section vise ainsi à étudier les écarts entre les coefficients de variation Merz & Wüthrich et Bootstrap, sur notre échantillon de données (cf. section 3.1). Le tableau suivant présente les statistiques obtenues sur la distribution des écarts relatifs :

Statistique	Écart relatif
Moyenne	960%
1er décile	-8%
1er quartile	-4%
Médiane	-1%
3-ième quartile	11%
9ième décile	41%

FIGURE 3.21 – Écart relatif des coefficients de variation Merz & Wüthrich et Bootstrap

Les écarts apparaissent réduits dans l'ensemble, avec tout de même des écarts très élevés comme en témoigne le niveau moyen ou le niveau du 9ième décile. En analysant les écarts de plus près, et en définissant un écart acceptable entre -10% et 10%, il ressort que 32% de l'échantillon présente un écart plus important. La moyenne ressort particulièrement élevée du fait d'écarts importants sur certains triangles. En retirant les plus gros écarts, la moyenne passerait à 189%. D'ailleurs, les écarts élevés feront l'objet d'un examen plus approfondi, conformément à l'algorithme présenté ci-dessous (cf. section 3.3).

Pour la suite de ce chapitre, nous nous intéressons à la méthode Bootstrap, et plus précisément à la convergence ainsi qu'à l'influence de la loi de simulation sur le résultat. Pour ces tests, seules les données présentant un écart faible entre les coefficients de variation Merz & Wüthrich et Bootstrap seront retenues dans l'échantillon, soit 68% de l'échantillon initial.

Remarque 3.5.1 Nous évoquons certaines fois les écarts entre les écarts-types obtenus par la méthode de Merz & Wüthrich et par méthode Bootstrap, et d'autres fois les écarts entre les coefficients de variation Merz & Wüthrich et Bootstrap. Ces deux grandeurs sont similaires et les deux expressions pourront être utilisées par la suite.

3.7 CONVERGENCE DANS LA MÉTHODE BOOTSTRAP

Les résultats précédents pour la méthode Bootstrap ont été établis en considérant un nombre de simulations $N = 10.000$. Ce choix est expliqué par des raisons de temps de calcul. Il est cependant légitime de soulever la question de savoir si ce nombre de simulations est suffisant. Nous avons vérifié empirique-

ment en considérant cette fois-ci $N = 100.000$ simulations.

La différence entre les deux résultats (10.000 et 100.000 simulations) est ensuite étudiée. Le tableau suivant présente les statistiques obtenues sur la distribution des écarts relatifs sur l'échantillon considéré (cf. section 3.6) :

Statistique	Écart relatif
Moyenne	0,11%
1er décile	-1,00%
1er quartile	-0,48%
Médiane	0,04%
3-ième quartile	0,47%
9ième décile	0,93%

FIGURE 3.22 – *Écart relatif des coefficients de variation Bootstrap avec 10.000 et 100.000 simulations*

Il ressort des écarts faibles entre les deux coefficients de variation, ce qui traduit une bonne convergence même avec un nombre de simulations égal à 10.000. Pour la suite, nous allons donc continuer à travailler sur des résultats Bootstrap obtenus à partir de 10.000 simulations.

Remarque 3.6.1 En considérant l'échantillon initial global, nous remarquons que certains écarts peuvent ressortir très élevés. Ces écarts importants peuvent s'expliquer par des erreurs dans les triangles, ce qui génère des résidus assez atypiques. Ainsi, en simulant 100.000 fois, ces résidus ont une probabilité plus importante d'être tirés, ce qui explique un écart-type final important. Néanmoins, ces cas sont également capturés par l'indicateur consistant à regarder la différence entre les coefficients de variation Merz & Wüthrich et Bootstrap, et n'apparaissent donc pas dans l'échantillon considéré ici.

3.8 MÉTHODE BOOTSTRAP AVEC D'AUTRES LOIS DE SIMULATION

Une des étapes de la méthode Bootstrap consiste à simuler une diagonale supplémentaire selon une certaine loi paramétrique de moyenne $C_{i,n+1-1}\lambda_{n+1-i}^k$ et de variance $C_{i,n+1-i}\hat{\sigma}_{n+1-i}$. Jusqu'à maintenant, les résultats pour la méthode Bootstrap ont été obtenus par simulation suivant une loi normale. Considérons maintenant d'autres lois de simulation, afin de comparer la matérialité.

Dans cette étude, nous nous sommes intéressés à trois différentes lois pour la méthode Bootstrap : la loi normale, la loi log-normale ainsi que la loi gamma.

Résultats obtenus à partir d'une loi log-normale

Une variable aléatoire X est dite suivre une loi log-normale si la variable $Y = \ln(X)$ suit une loi normale. Cette loi permet de modéliser des valeurs positives, et est souvent utilisée dans le cadre d'analyse de fiabilité et dans le domaine de la finance (par exemple pour modéliser le comportement d'actions). La loi log-normale est une loi de distribution continue définie par deux paramètres : μ et

σ^2 , son espérance et sa variance sont données par :

$$E[X] = e^{\mu + \frac{\sigma^2}{2}}$$

et $Var[X] = (e^{\sigma^2} - 1)e^{2\mu + \sigma^2}$.

Dans notre cas, l'espérance et la variance sont connues. Les relations équivalentes permettent alors d'obtenir les deux paramètres μ et σ^2 étant données ces deux premiers moments :

$$\mu = \ln(E[X]) - \frac{1}{2} \ln\left(1 + \frac{Var[X]}{(E[X])^2}\right)$$

et

$$\sigma^2 = \ln\left(1 + \frac{Var[X]}{(E[X])^2}\right).$$

Nous avons étudié la différence entre l'écart-type obtenu par la méthode Bootstrap avec simulation selon une loi normale, et l'écart-type obtenu par la méthode Bootstrap avec simulation selon une loi log-normale. Le tableau suivant présente les statistiques obtenus sur la distribution des écarts relatifs sur l'échantillon considéré (cf. section 3.6) :

Statistique	Écart relatif
Moyenne	-0,09%
1er décile	-0,71%
1er quartile	-0,31%
Médiane	-0,02%
3-ième quartile	0,31%
9ième décile	0,65%

FIGURE 3.23 – Écart relatif des coefficients de variation Bootstrap avec loi normale et loi log-normale

Il en ressort des écarts faibles entre les deux coefficients de variation, ce qui laisse à penser que remplacer la loi normale par une loi log-normale n'a que peu d'influence sur l'écart-type final. Intéressons-nous maintenant à la loi gamma, qu'en est-il pour cette loi ?

Résultats obtenus à partir d'une loi gamma

La loi gamma permet de modéliser des valeurs positives qui sont approximativement asymétriques vers la droite et supérieures à 0. Cette loi est à distribution continue définie par deux paramètres : α et β , sa densité, son espérance et sa variance sont données par :

$$f(x, \alpha, \beta) = \frac{x^{\alpha-1} e^{-\frac{x}{\beta}}}{\Gamma(\alpha) \beta^\alpha}$$

et

$$E[X] = \alpha\beta$$

et

$$Var[X] = \alpha\beta^2$$

Dans notre cas, l'espérance et la variance sont connues. Les relations équivalentes permettant alors d'obtenir les deux paramètres α et β sont données ces deux premiers moments :

$$\alpha = \frac{E[X]^2}{Var[X]}$$

et

$$\beta = \frac{Var[X]}{E[X]}.$$

Nous avons étudié la différence entre l'écart-type obtenu par la méthode Bootstrap avec simulation selon une loi normale, et l'écart-type obtenu par la méthode Bootstrap avec simulation selon un loi log-normale. Le tableau suivant présente les statistiques obtenus sur la distribution des écarts relatifs sur l'échantillon considéré (cf. section 3.6) :

Statistique	Écart relatif
Moyenne	-0,12%
1er décile	-0,79%
1er quartile	-0,37%
Médiane	-0,03%
3-ième quartile décile	0,33%
9ième décile	0,65%

FIGURE 3.24 – Écart relatif des coefficients de variation Bootstrap avec loi normale et loi gamma

Comme pour la loi log-normale, il ressort des écarts faibles pour l'échantillon considéré. Ces deux résultats (simulation par loi log-normale et par loi gamma) laissent donc à penser que la loi considérée n'a que peu d'influence sur l'écart-type obtenu en sortie de la méthode Bootstrap. Pour la suite, nous allons donc conserver une simulation par loi normale.

Remarque 3.7.1 Les écarts présentés ci-dessous ont été obtenus en considérant 82% de l'échantillon considéré, soit 56% de l'échantillon initial, du fait d'erreurs obtenues en simulant suivant la loi gamma.

Résultats numériques

4.1	Segment “autre assurance des véhicules à moteur”	54
4.2	Segment “responsabilité civile automobile”	57
4.3	Segment “responsabilité civile générale”	61
4.4	Segment “incendie et autres dommages aux biens”	64
4.5	Segment “pertes pécuniaires diverses”	67
4.6	Segment “protection juridique”	70
4.7	Segment “frais médicaux”	73
4.8	Segment “protection du revenu”	75
4.9	Segment “assistance”	79

Ce chapitre présente les résultats numériques obtenus pour les différents segments d'activité. Pour rappel, les "triangles" utilisés sont ceux des paiements, bruts et nets de réassurance. Les "triangles" de charges posent des problèmes de qualité de données comme indiqué dans la section 3.5. Les résultats entre les "triangles" de paiements bruts et nets sont très proches pour l'ensemble des segments d'activité étudiés (conformément aux résultats numériques présentés dans ce chapitre). Quand les résultats numériques pour les deux types de "triangles" ne sont pas présentés, le choix par défaut est le "triangle" de paiements bruts de réassurance.

4.1 SEGMENT "AUTRE ASSURANCE DES VÉHICULES À MOTEUR"

Ce segment d'activité regroupe, comme indiqué dans la section 3.1, les engagements d'assurance couvrant tout dommage subi par les véhicules à moteur terrestres. C'est en conséquence un segment d'activité à déroulement court dans le sens où les paiements de l'assureur se font relativement vite. Dans le cas d'un accident impliquant plusieurs véhicules à moteur, des recours peuvent arriver plus tardivement, mais cela reste globalement négligeable grâce à la convention IRSA (Indemnisation directe de l'assuré et de Recours entre Sociétés d'Assurance automobile). Signée par la plupart des sociétés d'assurance en France, elle est destinée à faciliter l'indemnisation des dommages matériels en cas d'accident de la circulation.

Taille de l'échantillon

Le tableau ci-dessous détaille le nombre de sociétés ("triangles") utilisées lors de l'étude de ce segment après le retraitement de données (cf. section 3.3).

Segment d'activité	Autre assurance des véhicules à moteur
Nr initial d'organismes	69
Nr organismes hors échantillon	8
Nr final	61
<i>Dont retraitements</i>	9

FIGURE 4.1 – Nombre de triangles analysés pour le segment d'activité "Autre assurance des véhicules à moteur"

La taille moyenne des triangles utilisés (après retraitements de données) est de 13 ans dont plus de 2/3 ont une longueur de 15 ans (la taille maximale renseignée dans le QRT). Pour un segment d'activité dont la liquidation se fait presque entièrement en 2 ans, une moyenne de 13 ans d'historique apparaît comme une taille fiable.

Duration des paiements

De manière approximative, la durée¹ de ce segment d'activité est d'environ 1,25 correspondant à 75% des paiements la première année et les 25% restant la deuxième année (en termes de durée $1,25 = 1 \times 75\% + 2 \times 25\%$).

1. Pour rappel, la formule de la durée est donnée dans la section 3.2 est cela revient à pondérer les paiements des sinistres par les années de paiements.

Des recours peuvent arriver au-delà de la deuxième année (et encore plus rarement des paiements), mais cela reste de second ordre. Sur l'échantillon final (i.e. après exclusions), le tableau 4.2 montre l'homogénéité des durations au sein de ce segment d'activité, que ce soit sur les paiements bruts ou les paiements nets.

Statistique	Triangle paiements bruts	Triangles de paiements nets
Moyenne	1,27	1,27
1er décile	1,12	0,12
1er quartile	1,14	0,14
3-ième quartile décile	1,30	1,31
9ième décile	1,38	1,40

FIGURE 4.2 – Statistiques sur la durée du segment "autre assurance des véhicules à moteur"

Facteur d'émergence

Pour rappel, le facteur d'émergence, défini comme le rapport entre le risque à un an (obtenu via la méthode de méthode de Merz & Wüthrich) et le risque à l'ultime (obtenu via la méthode Mack), permet de comprendre quelle partie du risque à l'ultime "émerge" pendant la première année. Sous l'hypothèse que l'intégralité des paiements d'une année de survenance est effectuée en 2 ans, le montant des réserves totales correspond aux réserves de la dernière année de survenance (i.e. $R = \sum_{i=2}^n R_i = R_n$) et il sera payé intégralement pendant l'année suivante. En conséquence, le risque à l'ultime serait la même chose que le risque à un an. Dit autrement, le facteur d'émergence est égal à 100%. En réalité, tout n'est pas réglé en 2 ans, mais le montant des règlements (positifs ou négatifs en cas de recours) reste globalement faible ce qui conduit à des facteurs d'émergence proches de 100% comme le montre le tableau 4.3.

Statistique	Triangle paiements bruts	Triangles de paiements nets
Moyenne	94%	93%
1er décile	82%	83%
1er quartile	92%	89%
3ème quartile	99%	99%
9ème décile	100%	100%

FIGURE 4.3 – Statistiques sur le facteur d'émergence du segment "autre assurance des véhicules à moteur"

Écart entre la méthode de Merz & Wüthrich et sa variante simplifiée (méthode USP)

Comme indiqué dans la section 2.9, une des deux méthodologies pour calculer les USP pour le risque de réserve repose sur une variante du modèle de Merz & Wüthrich (dite variante simplifiée). Pour rappel, en introduisant un développement de premier ordre, les auteurs ont simplifié la formule initiale (dit modèle "complet"). Cette approximation est "mathématiquement" à la baisse. Le tableau 4.4 ci-dessous montre que pour le cas du segment "autre assurance des véhicules à moteur" les écarts restent faibles.

Statistique	Triangle paiements bruts	Triangles de paiements nets
Moyenne	2%	2%
1er décile	0%	0%
1er quartile	0%	0%
3ème quartile	2%	3%
9ème décile	5%	5%

FIGURE 4.4 – Statistiques sur l'écart entre la méthode de Merz & Wüthrich et sa variante simplifiée (méthode USP) du segment "Autre assurance des véhicules à moteur"

Décroissance du coefficient de variation par rapport au volume

Afin d'estimer la décroissance du coefficient de variation (par rapport au modèle de Merz & Wüthrich) en fonction du volume en supposant une décroissance en loi puissance, nous représentons les données dans un repère "log-log" 4.5 transformant ainsi une fonction en puissance en droite. Le graphe 4.5 indique le résultat de cette régression linéaire, la pente est de $-0,19$ et le coefficient de détermination est de $R^2 = 0,39$. Nous avons testé également une décroissance logarithmique mais le coefficient de détermination est légèrement plus faible dans ce cas, $R^2 = 0,37$.

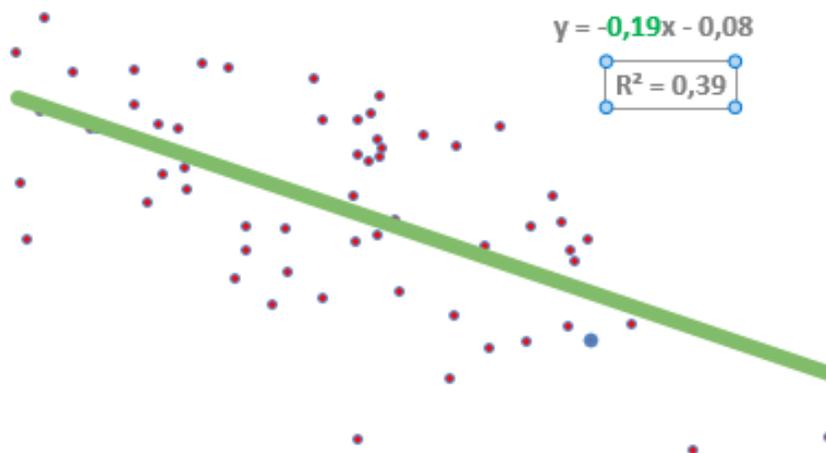


FIGURE 4.5 – Estimation du degré de la décroissance puissance sur le segment "autre assurance des véhicules à moteur"

Le graphe 4.6 compare les résultats obtenus précédemment avec le niveau de la formule standard. Un point se détache par son volume plus important mais influence que peu *in fine* le résultat de la régression (i.e. l'exposant de la fonction puissance).

De manière générale, le choc de la formule standard reste globalement en dessous des coefficients de variation du marché français. Quant au niveau du coefficient "représentatif" du marché français (qui serait calculé avec le modèle développé par EIOPA (cf. section 2.10, équation 2.10), il ressort fortement dépendant des deux derniers points à cause de leur volume. Si nous suivons le formalisme de l'EIOPA consistant à ne pas tenir compte de ces points, le résultat serait de 11,2%. Si nous tenons compte de ces points alors le coefficient serait

plutôt de 7,3%. Cet écart s’explique par la forte dépendance en volume : en effet la formule citée *supra* revient à faire une moyenne pondérée des coefficients de variation individuel par le carrée des volumes.

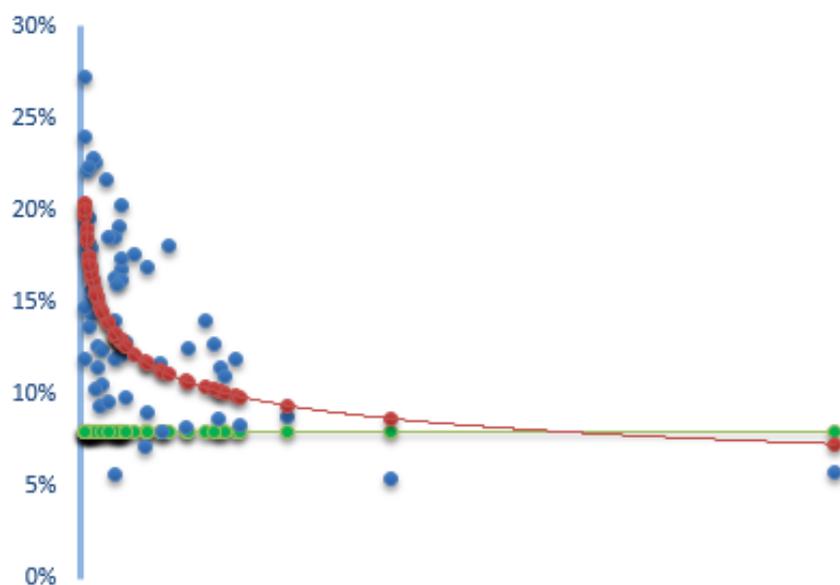


FIGURE 4.6 – Comparaison avec la formule standard pour le segment “autre assurance des véhicules à moteur”

Un montant intéressant est l’intersection entre le modèle théorique en loi puissance (courbe rouge) et le niveau de la formule standard (ligne verte). Ce montant indique de manière approximative pour quel volume la formule standard est représentative. Pour le segment “responsabilité civile automobile”, ce niveau est d’environ 50M. Comme indiqué dans le chapitre précédent, le niveau du BE de réserve intègre des frais supplémentaires, donc cela ne veut pas dire que les sociétés ayant un BE de réserve supérieur à 50M seraient désavantagées par la formule standard. En séparant l’ensemble des sociétés du marché par rapport à ce volume et dans le cadre de ce modèle théorique, 75% du marché français devrait être désavantagé par la formule standard et 25% devrait être avantagé. Bien évidemment, ce n’est pas la même proportion en termes de nombre de sociétés, c’est plutôt le contraire.

4.2 SEGMENT “RESPONSABILITÉ CIVILE AUTOMOBILE”

Ce segment d’activité regroupe, comme indiqué dans la section 3.1, les engagements d’assurance couvrant toute responsabilité découlant de l’utilisation d’un véhicule à moteur terrestre (y compris la responsabilité du transporteur). Comme tout segment de responsabilité civile, c’est un segment à déroulement long.

Taille de l’échantillon

Le tableau ci-dessous détaille le nombre de sociétés (“triangles”) utilisées lors de l’étude de ce segment après le retraitement de données (cf. section 3.3).

Il est à noter que seulement 2 sociétés ont été exclues à cause de la mauvaise qualité de données. Les 6 autres ont été exclues de l'échantillon de l'étude afin de garder un historique minimum de 9 ans à cause de la liquidation lente de cette garantie.

Segment d'activité	"responsabilité civile automobile"
Nr initial d'organismes	70
Nr organismes hors échantillon	8
Nr final	62
<i>Dont retraitements</i>	3

FIGURE 4.7 – Nombre de triangles analysés pour le segment d'activité "responsabilité civile automobile"

La taille moyenne des triangles utilisés est de 14 ans avec une très large majorité ayant une longueur de 15 ans (la taille maximale renseignée dans le QRT).

Duration des paiements

La duration des paiements est plus grande et bien plus dispersée que celle du segment d'activité "Autre assurance des véhicules à moteur" (cf. tableau 4.8 ci-dessous). En moyenne, 35% des paiements ont lieu la première année, 25% la deuxième, environ 7,5% les deux années d'après, 5% les deux années encore d'après, et pour le reste une bonne approximation pourrait être 2,5% pour les 6 années qui suivent. C'est ainsi une branche à liquidation lente.

Statistique	Triangle paiements bruts	Triangles de paiements nets
Moyenne	3,27	3,18
1er décile	2,72	2,23
1er quartile	2,86	2,73
3ème quartile	3,53	3,49
9ème décile	4,18	4,16

FIGURE 4.8 – Statistiques sur la duration du segment "responsabilité civile automobile"

Facteur d'émergence

Le facteur d'émergence de ce segment est plus petit que celui du segment "autre assurance des véhicules à moteur" car la liquidation est plus lente et donc une proportion plus faible du risque à l'ultime "émerge" à horizon d'un an.

Statistique	Triangle paiements bruts	Triangles de paiements nets
Moyenne	67%	70%
1er décile	60%	62%
1er quartile	63%	63%
3ème quartile	71%	74%
9ème décile	74%	83%

FIGURE 4.9 – Statistiques sur le facteur d'émergence du segment "responsabilité civile automobile"

Écart entre la méthode de Merz & Wüthrich et sa variante simplifiée (méthode USP)

Le tableau 4.10 ci-dessous montre que pour le cas du segment “responsabilité civile automobile” les écarts entre la variante simplifiée de Merz & Wüthrich et celle “complète” sont loin d’être négligeables. En moyenne, que ce soit sur le triangle des paiements bruts ou nets, le coefficient de variation calculé avec la variante simplifiée de Merz & Wüthrich est inférieur d’environ 15% par rapport à la méthode “complète”. En termes d’exigence en capital, un écart de cet ordre de grandeur devient non-négligeable.

Statistique	Triangle paiements bruts	Triangles de paiements nets
Moyenne	14%	16%
1er décile	11%	8%
1er quartile	14%	12%
3ème quartile	18%	18%
9ème décile	20%	20%

FIGURE 4.10 – Statistiques sur l’écart entre la méthode de Merz & Wüthrich et sa variante simplifiée (méthode USP) du segment “responsabilité civile automobile”

Décroissance du coefficient de variation par rapport au volume

Le graphe 4.11 indique le résultat de la régression linéaire relative au modèle de décroissance en loi puissance. L’exposant est de $-0,36$ et le coefficient de détermination est de $R^2 = 0,76$. Nous avons testé également une décroissance logarithmique mais le coefficient de détermination est plus faible dans ce cas, $R^2 = 0,57$.

Nous observons que la pente de la droite (ce qui correspond à l’exposant de la loi puissance) se rapproche davantage du modèle “jouet” avec un exposant de $-0,5$ et que la qualité de la régression est bien meilleure que celle du segment “autre assurance des véhicules à moteur”. Le segment de “responsabilité civile automobile” est celui où la qualité de la régression apparaît la meilleure et nous l’expliquons par l’homogénéité des risques. En effet, ce segment d’activité contient une seule garantie et qui plus est, est une garantie obligatoire. Cela rend la ligne d’activité homogène au sein d’une société mais aussi entre les sociétés du marchais français.

Le graphe 4.12 ci-dessous compare les résultats obtenus précédemment avec le niveau de la formule standard. Quelques points se détachent par leur volume plus important, mais ne changent pas la qualité de la régression. Le choc de la formule standard pénalise un certain nombre d’acteurs qui ont de volumes plus importants mais nous observons un nuage significatif de points autour du niveau du choc calibré par EIOPA. Concernant le coefficient “représentatif” du marché français (calculé avec le modèle développé par EIOPA (cf. section 2.10, équation 2.10), il est pour ce segment de risque fortement dépendant des gros volumes. A titre indicatif, avec l’ensemble des points, le résultat est de 4,4%, alors qu’en suivant le formalisme d’EIOPA consistant à ne pas tenir compte des gros volumes (dans notre cas, les sociétés avec un BE de réserve supérieur à 0,5G€), le résultat est de 8,8%.

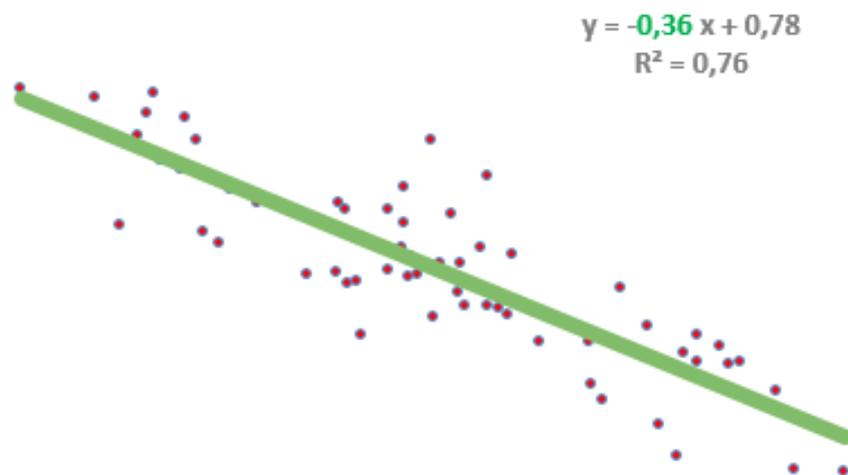


FIGURE 4.11 – Estimation du degré de la décroissance puissance sur le segment “responsabilité civile automobile”

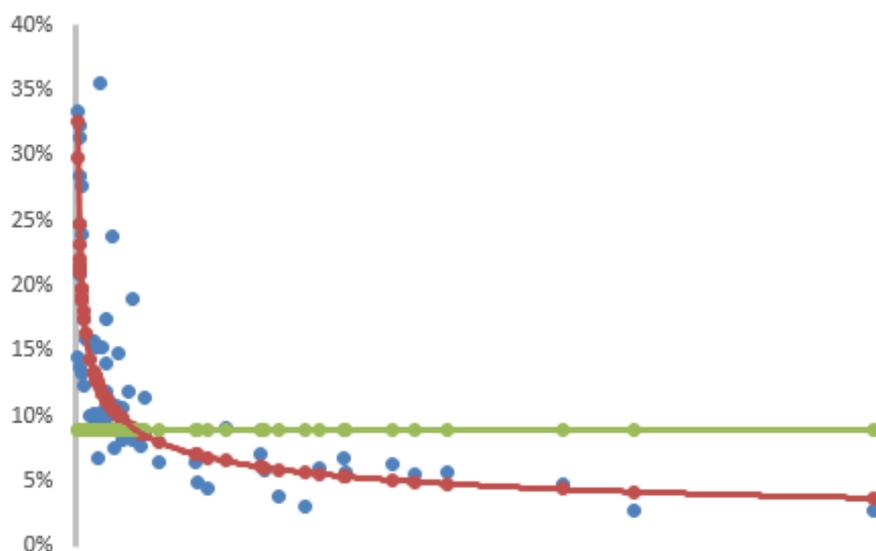


FIGURE 4.12 – Comparaison avec la formule standard pour le segment “responsabilité civile automobile”

L’intersection entre la loi puissance (courbe rouge) et le niveau de la formule standard (ligne verte) correspond à un volume d’environ 150M€. En séparant l’ensemble des sociétés du marché par rapport à ce volume et dans le cadre de ce modèle théorique, 80% du marché français devrait être désavantagé par la formule standard et 20% devrait être avantagé (en termes de volume de BE de réserve).

4.3 SEGMENT “RESPONSABILITÉ CIVILE GÉNÉRALE”

Ce segment d’activité regroupe, comme indiqué dans la section 3.1, les engagements d’assurance couvrant tous les domaines de la responsabilité à l’exception de ceux des segments “responsabilité civile automobile” et “assurance maritime, aérienne et transport”. Comme tout segment de responsabilité civile, c’est un segment à déroulement long.

Taille de l’échantillon

Le tableau ci-dessous détaille le nombre de sociétés (“triangles”) utilisées lors de l’étude de ce segment après le retraitement de données (cf. section 3.3). Il est à noter que seulement 3 sociétés ont été enlevées de l’échantillon à cause de la mauvaise qualité de données. Les 27 autres ont été enlevées de l’échantillon pour deux raisons :

- 5 sociétés à cause d’une longueur d’historique faible, comme pour le segment d’activité “responsabilité civile automobile” nous avons gardé des triangles de taille supérieure ou égale à 9 ans ;
- 22 sociétés ont été enlevée à cause la durée très importante des engagements sur ce segment de “responsabilité civile générale”. En effet, les engagements de responsabilité civile décennale en assurance construction et la responsabilité civile médicale sont des garanties qui ne sont pas homogènes avec le reste des garanties de responsabilité civile en termes de durée des engagements. De plus, en termes de paiements, un triangle de taille 15 ne serait pas suffisant pour prendre en compte la globalité des paiements pour une garantie comme celle de la responsabilité civile construction.

Segment d’activité	“responsabilité civile générale”
Nr initial d’organismes	75
Nr organismes hors échantillon	30
Nr final	45
<i>Dont retraitements</i>	4

FIGURE 4.13 – Nombre de triangles analysés pour le segment d’activité “responsabilité civile générale”

La taille moyenne des triangles utilisés est de 14 ans dont une très large majorité ayant une longueur de 15 ans (la taille maximale renseignée dans le QRT).

Duration des paiements

La durée des paiements est légèrement plus importante que celle du segment de “responsabilité civile automobile”, d’environ 0,5 sur l’ensemble de la distribution. En moyenne, 20% des paiements ont lieu la première année, 25% la deuxième, environ 10% les deux années d’après, 7,5% les deux années encore d’après, et pour le reste une bonne approximation pourrait être 3,33% pour les 6 années qui suivent. C’est également une branche à liquidation lente. La durée plus importante que celle du segment de “responsabilité civile automobile” s’explique principalement par des paiements moins importants la première année. En effet, pour le segment de “responsabilité civile automobile”,

les paiements liés aux dégâts matériels ont globalement lieu la première année.

A titre de comparaison, la durée moyenne des 22 sociétés laissées en dehors de l'échantillon (notamment à cause de leur durée plus importante) est de 5 ans.

Statistique	Triangle paiements bruts	Triangles de paiements nets
Moyenne	3,79	3,72
1er décile	2,90	2,88
1er quartile	3,32	3,32
3ème quartile	4,36	4,28
9ème décile	4,61	4,63

FIGURE 4.14 – Statistiques sur la durée du segment “responsabilité civile générale”

Facteur d'émergence

Les facteurs d'émergence de ce segment sont très proches de ceux du segment de “responsabilité civile automobile” et comme pour les autres segments, les résultats présentent peu d'écart entre les triangles de paiements bruts et nets.

Statistique	Triangle paiements bruts	Triangles de paiements nets
Moyenne	71%	71%
1er décile	63%	63%
1er quartile	66%	66%
3ème quartile	74%	74%
9ème décile	80%	82%

FIGURE 4.15 – Statistiques sur le facteur d'émergence du segment “responsabilité civile générale”

Écart entre la méthode de Merz & Wüthrich et sa variante simplifiée (méthode USP)

Comme pour le segment de “responsabilité civile automobile”, la méthode de l'USP (Merz & Wüthrich variante simplifiée) sous-estime le risque par rapport à la méthode “complète”. En moyenne, que ce soit sur le triangle des paiements bruts ou nets, l'écart sur le coefficient de variation est d'environ 15%.

Statistique	Triangle paiements bruts	Triangles de paiements nets
Moyenne	14%	15%
1er décile	9%	9%
1er quartile	11%	11%
3ème quartile	18%	18%
9ème décile	19%	19%

FIGURE 4.16 – Statistiques sur l'écart entre la méthode de Merz & Wüthrich et sa variante simplifiée (méthode USP) du segment “responsabilité civile générale”

Décroissance du coefficient de variation par rapport au volume

Conformément au graphe 4.17 ci-dessous, l'exposant de la décroissance en loi puissance est de $-0,22$ et le coefficient de détermination est de $R^2 = 0,47$. Pour ce segment, la décroissance logarithmique est d'une qualité proche (coefficient de détermination est de $R^2 = 0,45$) mais toujours en dessous de la décroissance en loi puissance.

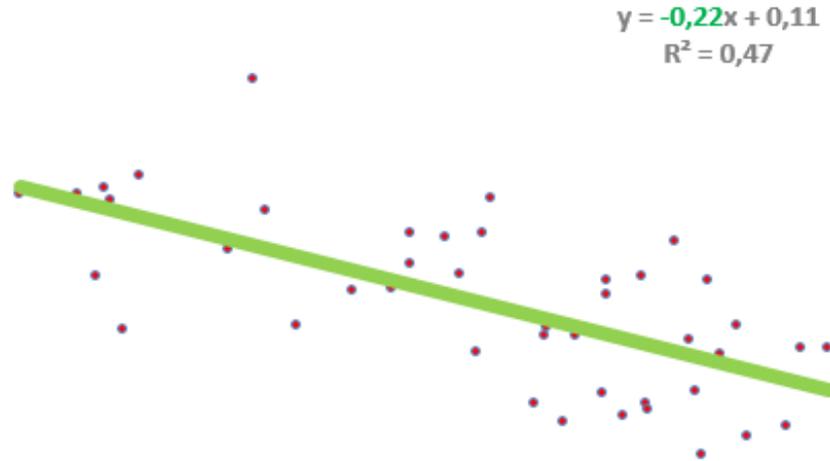


FIGURE 4.17 – Estimation du degré de la décroissance puissance sur le segment "responsabilité civile générale"

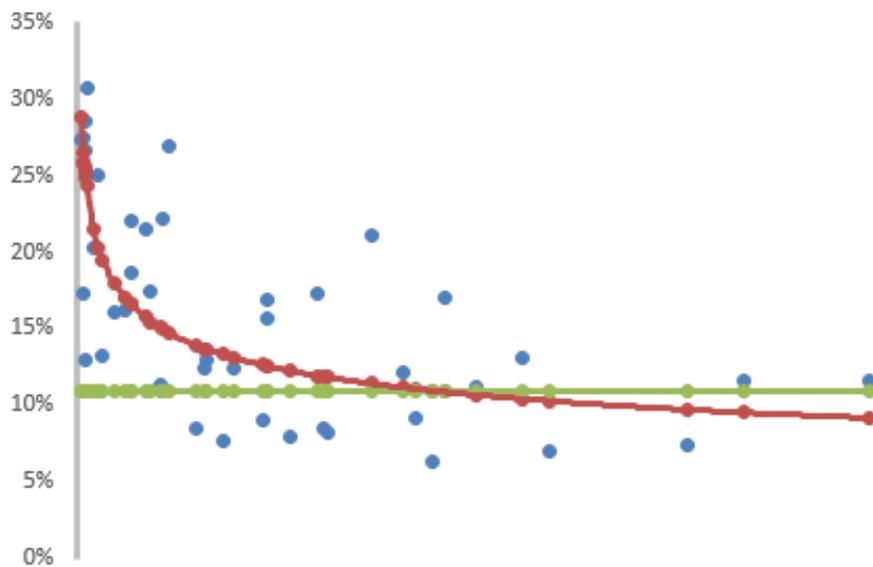


FIGURE 4.18 – Comparaison avec la formule standard pour le segment "responsabilité civile générale"

Le graphe ci-dessus compare les résultats obtenus précédemment avec le niveau de la formule standard. De manière générale, nous observons une très grande hétérogénéité des coefficients de variation, qui s'explique par la diversité des populations assurées. En effet, en plus de la responsabilité civile

“individuelle”, les responsabilités civiles professionnelles (ex. agriculteur, tuteur, service à la personne, etc.) comportent des risques différents et ne sont pas de la même nature que la responsabilité civile des entreprises. En termes de modélisation mathématique, les risques ne sont pas identiques et cela se traduit par ces résultats hétérogènes. Pour mieux étudier ce segment de risque, il faudrait faire des études par groupe de risques homogènes, ce qui n’est pas réalisable avec les données à disposition.

Le choc de la formule standard est globalement représentatif du marché français malgré. Le coefficient de variation “représentatif” du marché français (calculé avec le formalisme développé par EIOPA (cf. section 2.10, équation 2.10), est de 11,2%. Il est à noter que les “gros volumes” ne font pas partie de l’échantillon de l’étude.

Le volume correspondant à l’intersection de la loi puissance (courbe rouge) avec le niveau de la formule standard (courbe verte) est d’environ 50M€.

4.4 SEGMENT “INCENDIE ET AUTRES DOMMAGES AUX BIENS”

Ce segment d’activité regroupe, comme indiqué dans la section 3.1, les engagements d’assurance couvrant tout dommage subi par les biens (en dehors des segments “autre assurance des véhicules à moteur” et “maritime, aérienne et transport”), par suite d’un incendie, d’une explosion, des forces de la nature y compris les tempêtes, la grêle ou le gel, d’un accident nucléaire, d’un affaissement de terrain ou d’un autre événement tel que le vol. De manière plus simple, ce segment est appelé communément “dommages aux biens”.

Taille de l’échantillon

Le tableau ci-dessous détaille le nombre de sociétés utilisées lors de l’étude de ce segment après le retraitement de données (cf. section 3.3). Seulement 5 sociétés ont été enlevées à cause de la mauvaise qualité de données. Les 19 autres ont été enlevées de l’échantillon pour deux raisons :

- 5 sociétés à cause d’une longueur d’historique faible, nous avons gardé des triangles de taille supérieure ou égale à 7 ans ;
- 14 sociétés ont été enlevée à cause la durée très importante des engagements. En effet, les engagements de dommages ouvrages en assurance construction ne sont pas homogènes avec le reste des garanties du segment “incendie et autres dommages aux biens” à cause de la durée des engagements. De plus, comme pour la garantie de responsabilité civile décennale, un triangle de taille 15 ne serait pas suffisant pour prendre en compte la globalité des paiements.

La taille moyenne des triangles utilisés est de 13 ans avec une large majorité ayant une longueur de 15 ans (la taille maximale renseignée dans le QRT).

Duration des paiements

La durée des paiements est bien plus faible que les segments de “responsabilité civile automobile”, mais supérieur au segment “autre assurance des vé-

Segment d'activité	“incendie et autres dommages aux biens”
Nr initial d'organismes	76
Nr organismes hors échantillon	24
Nr final	52
<i>Dont retraitements</i>	3

FIGURE 4.19 – Nombre de triangles analysés pour le segment d'activité “incendie et autres dommages aux biens”

hicules à moteur” (d'environ 0,5 sur l'ensemble de la distribution). En moyenne, la liquidation totale des sinistres relatifs à une année de survenance prend 5 ans avec la cadence suivante : 60%, 30%, 5%, 2,5% et respectivement 2,5%.

Statistique	Triangle paiements bruts	Triangles de paiements nets
Moyenne	1,69	1,63
1er décile	1,49	1,39
1er quartile	1,54	1,49
3ème quartile	1,85	1,76
9ème décile	1,95	1,88

FIGURE 4.20 – Statistiques sur la duration du segment “incendie et autres dommages aux biens”

Facteur d'émergence

De manière inverse par rapport à la duration, les facteurs d'émergence de ce segment (en grande partie entre 80% et 90%), sont supérieurs à ceux des segments de responsabilité civile et inférieurs à ceux du segment “autre assurance des véhicules à moteur”.

Statistique	Triangle paiements bruts	Triangles de paiements nets
Moyenne	86%	87%
1er décile	75%	78%
1er quartile	80%	83%
3ème quartile	93%	93%
9ème décile	95%	95%

FIGURE 4.21 – Statistiques sur le facteur d'émergence du segment “incendie et autres dommages aux biens”

Écart entre la méthode de Merz & Wüthrich et sa variante simplifiée (méthode USP)

Comme pour les autres indicateurs, l'écart entre la méthode de l'USP (Merz & Wüthrich variante simplifiée) et la version “complète” (de 5% en moyenne) est inférieur aux écarts observés sur les segments de responsabilité civile (en moyenne 15%).

Statistique	Triangle paiements bruts	Triangles de paiements nets
Moyenne	5%	4%
1er décile	1%	1%
1er quartile	2%	2%
3ème quartile	7%	6%
9ème décile	9%	9%

FIGURE 4.22 – Statistiques sur l'écart entre la méthode de Merz & Wüthrich et sa variante simplifiée (méthode USP) du segment "incendie et autres dommages aux biens"

Décroissance du coefficient de variation par rapport au volume

Conformément au graphique 4.23 ci-dessous, l'exposant de la décroissance en loi puissance est de $-0,27$ et le coefficient de détermination est de $R^2 = 0,46$. Pour ce segment également, la décroissance logarithmique est de moins bonne qualité que celle en loi puissance, avec un coefficient de détermination $R^2 = 0,38$.

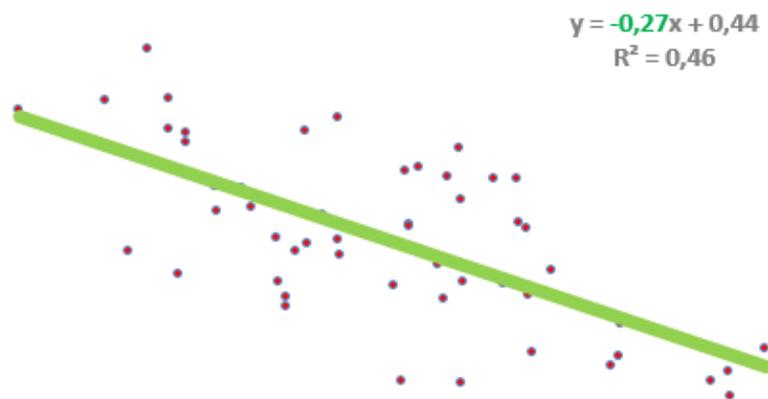


FIGURE 4.23 – Estimation du degré de la décroissance puissance sur le segment "incendie et autres dommages aux biens"

De manière générale, le choc de la formule standard reste globalement en dessous des coefficients de variation du marché français. Quant au niveau du coefficient "représentatif" du marché français (calculé avec le modèle développé par EIOPA (cf. section 2.10, équation 2.10), il est fortement dépendant des quatre derniers points à cause de leur volume. Si nous suivons le formalisme de l'EIOPA consistant à ne pas tenir compte de ces points, le résultat serait de 12,4%. Si nous tenons compte de ces points alors le coefficient serait plutôt de 8,6%.

Le volume correspondant à l'intersection de la loi puissance (courbe rouge) avec le niveau de la formule standard (courbe verte) est d'environ 175M€.

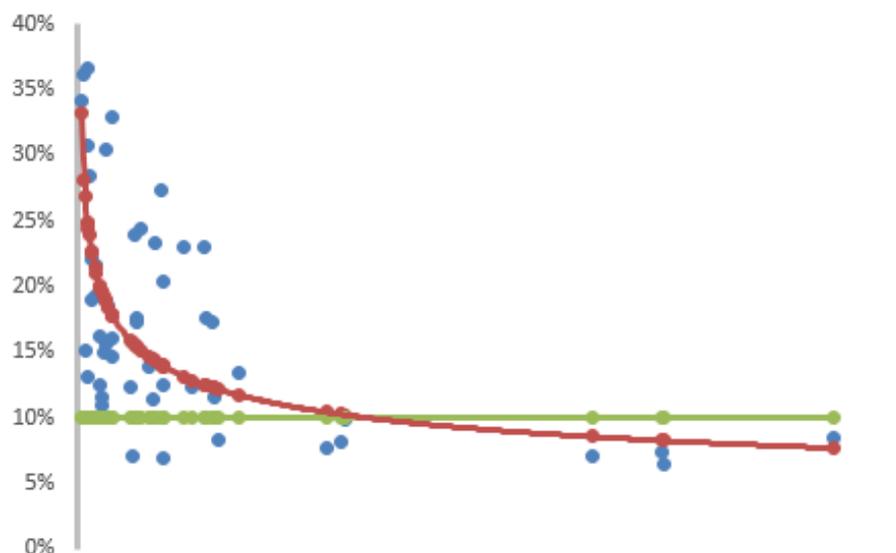


FIGURE 4.24 – Comparaison avec la formule standard pour le segment “incendie et autres dommages aux biens”

4.5 SEGMENT “PERTES PÉCUNIAIRES DIVERSES”

Ce segment d’activité regroupe, comme indiqué dans la section 3.1, les engagements d’assurance couvrant le risque d’emploi, l’insuffisance de recettes, les intempéries, la perte de bénéfices, la persistance de frais généraux, les frais commerciaux imprévus, la diminution de la valeur vénale, la perte de loyers ou de revenus, les pertes commerciales indirectes autres que celles mentionnées plus haut, les pertes pécuniaires non commerciales, ainsi que *tout autre risque d’assurance non-vie qui n’est pas couvert par les lignes d’activité déjà citées*. Ce segment couvre donc des risques qui ne sont pas homogènes. Il est ainsi attendu que cela se traduise dans les résultats numériques.

Taille de l’échantillon

Le tableau ci-dessous détaille le nombre de sociétés utilisées lors de l’étude de ce segment après le retraitement de données (cf. section 3.3). Sur les 20 sociétés exclus, 4 sociétés ont été enlevées à cause de la mauvaise qualité des données, les 15 autres ont été écartées à cause du volume très faible (inférieure à 0,5M€). Une dernière société a été enlevée à cause d’un volume bien trop grand par rapport aux autres.

Segment d’activité	“pertes pécuniaires diverses”
Nr initial d’organismes	46
Nr organismes hors échantillon	20
Nr final	26
<i>Dont retraitements</i>	0

FIGURE 4.25 – Nombre de triangles analysés pour le segment d’activité “pertes pécuniaires diverses”

La taille moyenne des triangles utilisés est de 11 ans dont peu avec une longueur de 15 ans (la taille maximale renseignée dans le QRT).

Duration des paiements

La duration des paiements est similaire à celle du segment “incendie et autres dommages aux biens” mais la distribution est plus dispersée (l'écart-type est 2 fois plus grand). En moyenne, la liquidation totale des sinistres relatifs à une année de survenance prend 5 ans avec la cadence suivante : 60%, 30%, 5% et les 5% restant les trois années qui suivent.

Statistique	Triangle paiements bruts	Triangles de paiements nets
Moyenne	1,76	1,72
1er décile	1,29	1,25
1er quartile	1,37	1,35
3ème quartile	2,01	1,97
9ème décile	2,56	2,36

FIGURE 4.26 – Statistiques sur la duration du segment “pertes pécuniaires diverses”

Facteur d'émergence

Les facteurs d'émergence sont également proches de ceux du segment “incendie et autres dommages aux biens” mais légèrement supérieurs.

Statistique	Triangle paiements bruts	Triangles de paiements nets
Moyenne	92%	92%
1er décile	71%	72%
1er quartile	92%	92%
3ème quartile	99%	99%
9ème décile	100%	100%

FIGURE 4.27 – Statistiques sur le facteur d'émergence du segment “pertes pécuniaires diverses”

Écart entre la méthode de Merz & Wüthrich et sa variante simplifiée (méthode USP)

Les écarts entre la méthode de l'USP (Merz & Wüthrich variante simplifiée) et celle “complète” sont également proches de ceux du segment “incendie et autres dommages aux biens” mais avec quelques cas où les écarts deviennent supérieurs à 15%.

Statistique	Triangle paiements bruts	Triangles de paiements nets
Moyenne	4%	4%
1er décile	0%	0%
1er quartile	1%	1%
3ème quartile	6%	5%
9ème décile	15%	14%

FIGURE 4.28 – Statistiques sur l'écart entre la méthode de Merz & Wüthrich et sa variante simplifiée (méthode USP) du segment “pertes pécuniaires diverses”

Décroissance du coefficient de variation par rapport au volume

Conformément au graphe 4.29 ci-dessous, l'exposant de la décroissance exponentielle est de $-0,21$ et le coefficient de détermination est de $R^2 = 0,45$. Pour ce segment également, la décroissance logarithmique est d'une moins bonne qualité que celle en loi puissance, le coefficient de détermination est de $R^2 = 0,36$.

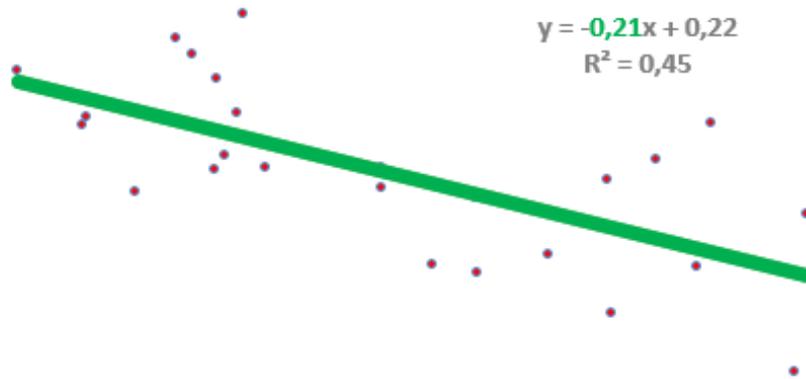


FIGURE 4.29 – Estimation du degré de la décroissance puissance sur le segment "pertes pécuniaires diverses"

Le graphe suivant compare les résultats obtenus précédemment avec le niveau de la formule standard. Les points ressortent très dispersés, expliqué par l'hétérogénéité forte des risques. Le coefficient de variation "représentatif" du marché français (calculé avec le formalisme développé par EIOPA (cf. section 2.10, équation 2.10) est identique à celui calibré par EIOPA, à savoir 20%.

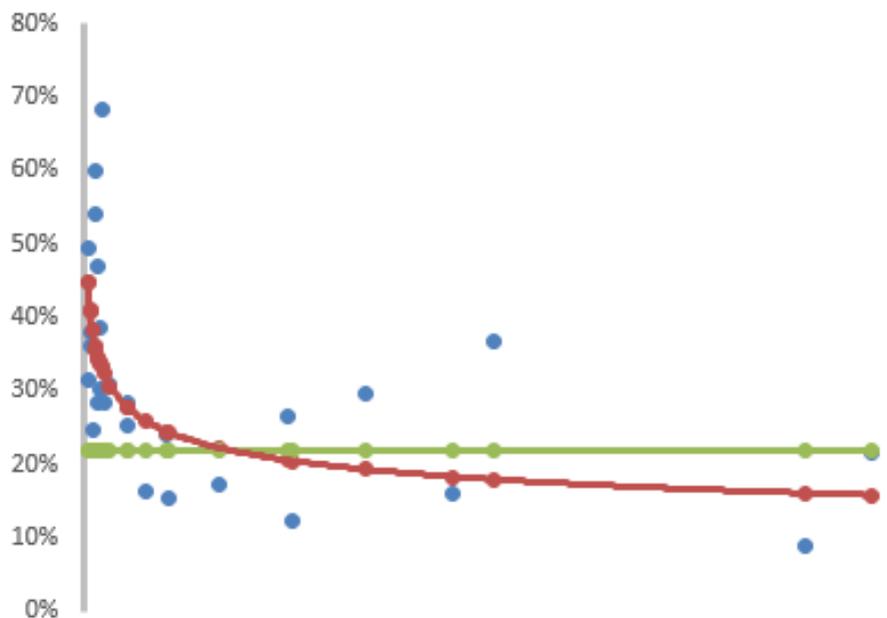


FIGURE 4.30 – Comparaison avec la formule standard pour le segment "pertes pécuniaires diverses"

Le volume correspondant à l'intersection de la loi puissance (courbe rouge)

avec le niveau de la formule standard (courbe verte) est d'environ 30M€. En séparant l'ensemble des sociétés du marché par rapport à ce volume et dans le cadre de ce modèle théorique, 75% du marché français devrait être désavantagé par la formule standard et 25% devrait être avantagé (en termes de volume de BE de réserve).

4.6 SEGMENT “PROTECTION JURIDIQUE”

Ce segment d'activité regroupe, comme indiqué dans la section 3.1, les engagements d'assurance couvrant les frais juridiques et le coût des actions en justice. Alors que ce segment de risque couvre une seule garantie, son homogénéité n'est pas garantie. En effet, ce type de garantie est proposé par les assureurs pour accompagner les garanties de responsabilité civile et de dommages (que ce soit pour le segment d'activité “incendie et autres dommages aux biens”, “maritime, aérienne et transport” ou “pertes pécuniaires diverses”. En fonction des contrats commercialisés par les différentes sociétés, le coût des actions en justice peut coûter plus ou moins cher et pour des périodes de temps différentes.

Taille de l'échantillon

Le tableau ci-dessous détaille le nombre de sociétés (“triangles”) utilisées lors de l'étude de ce segment après le retraitement de données (cf. section 3.3). Il est à noter que seulement 2 sociétés ont été enlevées à cause de la mauvaise qualité de données. Les 15 autres ont été enlevées de l'échantillon pour deux raisons :

- à cause d'un volume trop faible (dans ce cas inférieure à 0,2Me), 6 sociétés n'ont pas été prises en compte ;
- 9 sociétés ont été enlevées de l'échantillon final car la longueur d'historique était considérée insuffisante. Plus précisément, nous avons gardé les triangles avec une longueur supérieure ou égal à 8 ans.

Segment d'activité	“protection juridique”
Nr initial d'organismes	61
Nr organismes hors échantillon	17
Nr final	44
<i>Dont retraitements</i>	2

FIGURE 4.31 – Nombre de triangles analysés pour le segment d'activité “protection juridique”

La taille moyenne des triangles utilisés est de 13 ans avec une large majorité ayant une longueur de 15 ans (la taille maximale renseignée dans le QRT).

Duration des paiements

La duration des paiements est plus petite que celles des segments de responsabilité civile et supérieure à celles des segments courts (“assurance des véhicules à moteur”, “incendie et autres dommages aux biens” ou “pertes pécuniaires diverses”). En moyenne, la liquidation totale des sinistres relatifs à une année de survenance prend 7 ans avec la cadence suivante : 30%, 35%, 15%, 10%, 5%, 2, 5% et respectivement 2, 5% la septième année.

Statistique	Triangle paiements bruts	Triangles de paiements nets
Moyenne	2,46	2,39
1er décile	1,94	1,74
1er quartile	2,30	2,24
3ème quartile	2,69	2,69
9ème décile	2,87	2,94

FIGURE 4.32 – Statistiques sur la duration du segment “protection juridique”

Facteur d'émergence

Alors que la duration du segment de “protection juridique” est inférieure à celles des segments de responsabilité civile et supérieure à celles des segments courts (cf. *supra*), cela n'est pas le cas pour les facteurs d'émergence. Ils sont globalement semblables à ceux du segment “incendie et autres dommages aux biens”. Ce résultat est plus difficilement interprétable, mais probablement que le risque réside dans le temps de liquidation des sinistres (associé au temps de liquidation des sinistres litigieux) et moins dans leur coût ultime (qui est proportionnel au temps).

Statistique	Triangle paiements bruts	Triangles de paiements nets
Moyenne	85%	86%
1er décile	76%	77%
1er quartile	81%	80%
3ème quartile	91%	792%
9ème décile	93%	95%

FIGURE 4.33 – Statistiques sur le facteur d'émergence du segment “protection juridique”

Écart entre la méthode de Merz & Wüthrich et sa variante simplifiée (méthode USP)

Les écarts entre la méthode de l'USP (Merz & Wüthrich variante simplifiée) et la version “complète” reste aussi similaire à ceux du segment “d'incendie et autres dommages aux biens” mais légèrement plus faibles.

Statistique	Triangle paiements bruts	Triangles de paiements nets
Moyenne	6%	6%
1er décile	2%	2%
1er quartile	3%	3%
3ème quartile	8%	8%
9ème décile	11%	11%

FIGURE 4.34 – Statistiques sur l'écart entre la méthode de Merz & Wüthrich et sa variante simplifiée (méthode USP) du segment “protection juridique”

Décroissance du coefficient de variation par rapport au volume

Conformément au graphe 4.23 ci-dessous, l'exposant de la décroissance en loi puissance est de $-0,29$ et le coefficient de détermination est de $R^2 = 0,56$. Pour ce segment également, la décroissance logarithmique est d'une moins

bonne qualité que celle en loi puissance, le coefficient de détermination est de $R^2 = 0,36$.

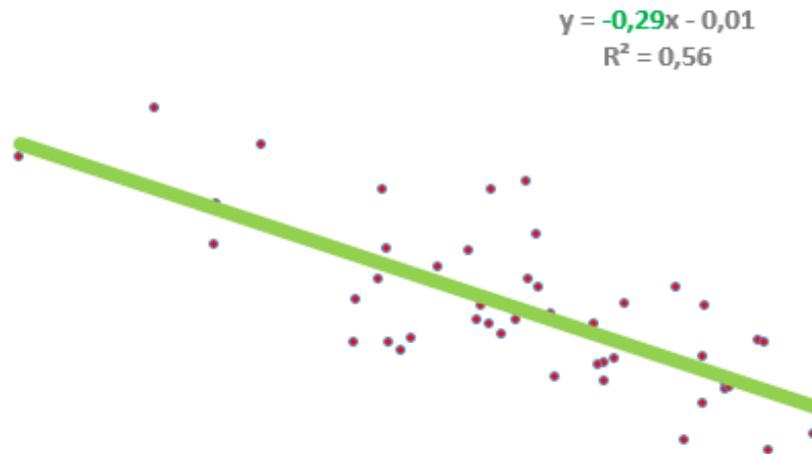


FIGURE 4.35 – Estimation du degré de la décroissance puissance sur le segment “protection juridique”

De manière générale, le choc de la formule standard reste supérieur aux coefficients de variation du marché français, mais les écarts restent faibles. Quant au niveau du coefficient “représentatif” du marché français (calculé avec le modèle développé par EIOPA (cf. section 2.10, équation 2.10), il est dépendant des quatre derniers points à cause de leur volume. Si nous suivons le formalisme de l’EIOPA consistant à ne pas tenir compte de ces points, le résultat serait de 5,8%. Si nous tenons compte de ces points, alors le résultat serait plutôt de 5%.

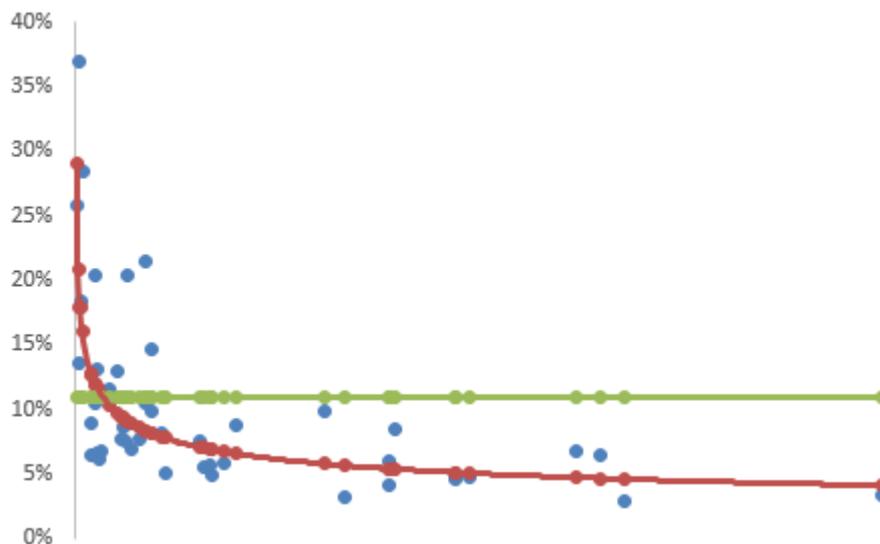


FIGURE 4.36 – Comparaison avec la formule standard pour le segment “protection juridique”

Le volume correspondant à l’intersection de la loi puissance (courbe rouge) avec le niveau de la formule standard (courbe verte) est d’environ 10M€. En séparant l’ensemble des sociétés du marché par rapport à ce volume et dans le

cadre de ce modèle théorique, 80% du marché français devrait être désavantagé par la formule standard et 20% devrait être avantagé (en termes de volume de BE de réserve).

4.7 SEGMENT “FRAIS MÉDICAUX”

Ce segment d’activité regroupe, comme indiqué dans la section 3.1, les engagements d’assurance couvrant de frais médicaux dans les cas où l’activité sous-jacente n’est pas exercée sur une base technique similaire à celle de l’assurance vie. C’est le segment d’activité avec la liquidation la plus rapide.

Taille de l’échantillon

Le tableau ci-dessous détaille le nombre de sociétés utilisées lors de l’étude de ce segment après le retraitement de données (cf. section 3.3). C’est le segment d’activité le plus représenté en termes de nombre de sociétés. La plupart des exclusions sont motivées par des problèmes de qualité de données (13 sociétés). 9 sociétés supplémentaires ont été enlevées de l’échantillon à cause du volume trop petit (inférieur à 0,2M€).

Segment d’activité	“frais médicaux”
Nr initial d’organismes	89
Nr organismes hors échantillon	22
Nr final	67
<i>Dont retraitements</i>	4

FIGURE 4.37 – Nombre de triangles analysés pour le segment d’activité “frais médicaux”

La taille moyenne des triangles utilisés est de 11 ans dont une très large majorité ayant une longueur de 15 ans (la taille maximale renseignée dans le QRT). Il est à noter qu’aucun filtre n’a été mis sur la longueur des triangles, la taille minimale étant de 4 ans.

Duration des paiements

La duration des paiements est légèrement inférieure au segment “autre assurance des véhicules à moteur” avec un écart-type de la distribution assez faible. En moyenne, 80% des paiements ont lieu la première année et 20% la deuxième année.

Statistique	Triangle paiements bruts	Triangles de paiements nets
Moyenne	1,20	1,20
1er décile	1,10	1,10
1er quartile	1,11	1,11
3ème quartile	1,22	1,22
9ème décile	1,32	1,30

FIGURE 4.38 – Statistiques sur la duration du segment “frais médicaux”

Facteur d'émergence

Les facteurs d'émergence de ce segment sont, comme ceux du segment de "autre assurance des véhicules à moteur", très proches de 100% car le risque à l'ultime coïncide avec le risque à un an.

Statistique	Triangle paiements bruts	Triangles de paiements nets
Moyenne	97%	98%
1er décile	92%	92%
1er quartile	97%	97%
3ème quartile	100%	100%
9ème décile	100%	100%

FIGURE 4.39 – Statistiques sur le facteur d'émergence du segment "frais médicaux"

Écart entre la méthode de Merz & Wüthrich et sa variante simplifiée (méthode USP)

Comme pour le segment de "autre assurance des véhicules à moteur", les écarts entre la méthode de l'USP (Merz & Wüthrich variante simplifiée) et la variante "complète" sont assez faibles.

Statistique	Triangle paiements bruts	Triangles de paiements nets
Moyenne	1%	1%
1er décile	0%	0%
1er quartile	0%	0%
3ème quartile	2%	2%
9ème décile	3%	3%

FIGURE 4.40 – Statistiques sur l'écart entre la méthode de Merz & Wüthrich et sa variante simplifiée (méthode USP) du segment "frais médicaux"

Décroissance du coefficient de variation par rapport au volume pour le segment "frais médicaux"

Conformément au graphe 4.23 ci-dessous, l'exposant de la décroissance en loi puissance est de $-0,18$ et le coefficient de détermination est de $R^2 = 0,24$. Pour ce segment également la décroissance logarithmique est d'une moins bonne qualité que celle en loi puissance, le coefficient de détermination est de $R^2 = 0,20$.

Le graphe suivant compare les résultats obtenus précédemment avec le niveau de la formule standard. Nous observons facilement une forte dispersion des points (ce qui explique aussi le niveau faible du coefficient de détermination). Nous n'arrivons pas à expliquer les raisons, sauf l'intuition d'une mauvaise qualité de données sur ce segment de risque. Le coefficient de variation "représentatif" du marché français (calculé avec le formalisme développé par EIOPA (cf. section 2.10, équation 2.10) est de 6,7%, supérieur à celui de la formule standard 5,7%. Cela s'observe sur le graphe ci-dessous qui montre que le choc de la formule standard est globalement en faveur de la plupart des sociétés du marché.

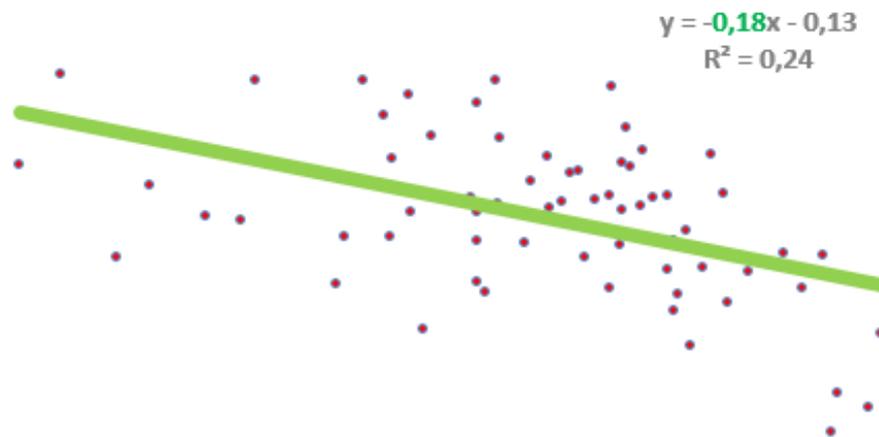


FIGURE 4.41 – Estimation du degré de la décroissance puissance

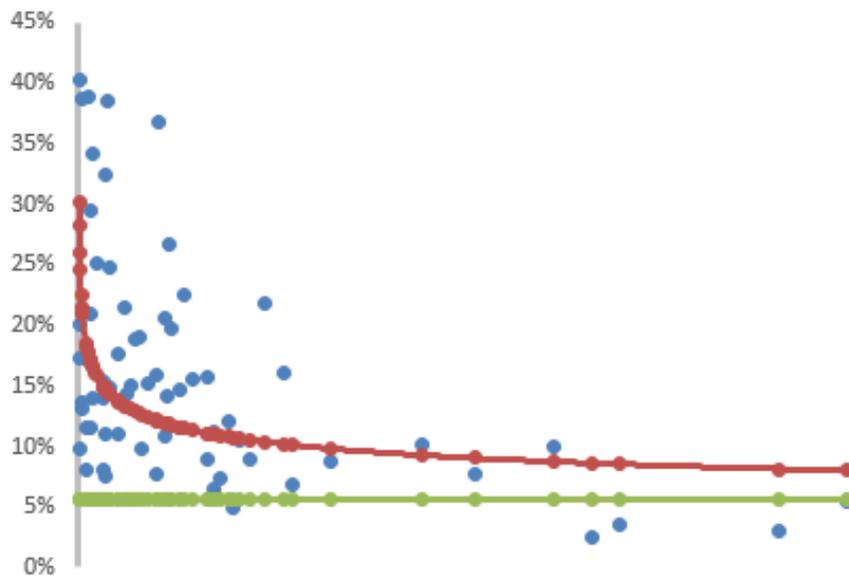


FIGURE 4.42 – Comparaison avec la formule standard pour le segment "frais médicaux"

Sur l'intervalle considéré dans le graphe ci-dessus, il n'y a pas d'intersection entre le modèle théorique et la droite représentant la formule standard, illustrant ainsi que cette dernière est globalement favorable aux marché français pour le segment de "frais médicaux".

4.8 SEGMENT "PROTECTION DU REVENU"

Ce segment d'activité regroupe, comme indiqué dans la section 3.1, les engagements d'assurance de protection du revenu dans les cas où l'activité sous-jacente n'est pas exercée sur une base technique similaire à celle de l'assurance vie. Les garanties proposées dans le cadre de ce segment comprennent des ga-

ranties liées à l'incapacité, ce qui explique que ce segment soit à déroulement plutôt long et avec des montants provisionnés pouvant être importants.

Taille de l'échantillon

Le tableau ci-dessous détaille le nombre de sociétés ("triangles") utilisées lors de l'étude de ce segment après le retraitement de données (cf. section 3.3). Il est à noter que seules 2 sociétés ont été enlevées à cause de la mauvaise qualité de données. Les 30 autres ont été enlevées de l'échantillon pour trois raisons distinctes :

- 21 sociétés à cause d'une longueur d'historique faible, le segment n'étant pas à déroulement court nous avons gardé des triangles de taille supérieure ou égale à 7 ans ;
- 8 sociétés ont été enlevées à cause d'un volume trop faible (inférieur à 1M€) ;
- 1 sociétés a été enlevée à cause d'un volume beaucoup plus élevé que les autres sociétés.

Segment d'activité	"protection du revenu"
Nr initial d'organismes	94
Nr organismes hors échantillon	32
Nr final	62
<i>Dont retraitements</i>	3

FIGURE 4.43 – Nombre de triangles analysés pour le segment d'activité "protection du revenu"

Cette activité de protection du revenu est exercée par l'ensemble des institutions de prévoyance, organismes spécifiques au marché français, ce qui explique le nombre important de sociétés reportés dans le tableau. La taille moyenne des triangles utilisés est de 13 ans, avec plus de la moitié ayant une longueur de 15 ans (la taille maximale renseignée dans le QRT).

Duration des paiements

La duration des paiements apparait relativement importante en comparaison à des segments très courts tels que "frais médicaux" ou "autre assurance des véhicules à moteur". Cependant, la duration pour ce segment reste tout de même inférieure à celles obtenues pour les segments de "responsabilité civile". En moyenne, la liquidation totale des sinistres relatifs à une année de survenance prend 7 ans avec la cadence suivante : 30%, 37,5%, 17,5%, 7,5% et les 7,5% restants repartis sur trois ans.

Statistique	Triangle paiements bruts	Triangles de paiements nets
Moyenne	2,50	2,46
1er décile	1,80	1,63
1er quartile	1,98	1,92
3ème quartile	2,74	2,74
9ème décile	3,80	3,82

FIGURE 4.44 – Statistiques sur la duration du segment "protection du revenu"

Plusieurs sociétés présentent une durée plus élevée, autour de 4,5 – 5. Ainsi, dans ce cas, le critère de durée n’a pas été utilisé pour sélectionner les données. En effet, le segment *protection du revenu* mélange deux types de risque, le risque emprunteur (de durée plus longue) et les garanties d’incapacité-invalidité.

Facteur d’émergence

Là encore, les facteurs d’émergence restent plus petits que ceux pour un segment à déroulement court, mais il ressort plus important que celui obtenu pour les segments “responsabilité civile”. Comme pour les autres segments, peu d’écart entre les triangles de paiements bruts et nets.

Statistique	Triangle paiements bruts	Triangles de paiements nets
Moyenne	85%	85%
1er décile	73%	72%
1er quartile	78%	77%
3ème quartile	94%	94%
9ème décile	98%	98%

FIGURE 4.45 – Statistiques sur le facteur d’émergence du segment “protection du revenu”

Écart entre la méthode de Merz & Wüthrich et sa variante simplifiée (méthode USP)

Les écarts entre les deux méthodes apparaissent non négligeables, mais restent tout de même inférieurs aux écarts obtenus pour les segments de “responsabilité civile”.

Statistique	Triangle paiements bruts	Triangles de paiements nets
Moyenne	6%	6%
1er décile	1%	1%
1er quartile	2%	2%
3ème quartile	10%	10%
9ème décile	12%	14%

FIGURE 4.46 – Statistiques sur l’écart entre la méthode de Merz & Wüthrich et sa variante simplifiée (méthode USP) du segment “protection du revenu”

Décroissance du coefficient de variation par rapport au volume pour le segment “protection du revenu”

Conformément au graphe 4.23 ci-dessous, l’exposant de la décroissance en loi puissance est de $-0,40$ et le coefficient de détermination est de $R^2 = 0,55$. Pour ce segment également, la décroissance logarithmique est d’une moins bonne qualité que celle en loi puissance, le coefficient de détermination est de $R^2 = 0,48$. Nous observons que la pente de la droite (ce qui correspond à l’exposant de la loi puissance) se rapproche davantage du modèle “jouet” avec un exposant de $-0,5$ et que la qualité de la régression apparaît bonne, ce qui traduit d’une certaine homogénéité au sein de ce segment.

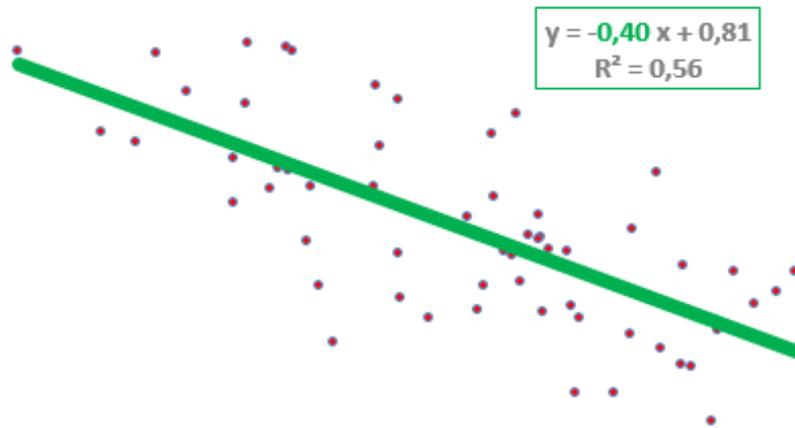


FIGURE 4.47 – Estimation du degré de la décroissance puissance

Le graphe suivant compare les résultats obtenus précédemment avec le niveau de la formule standard. Quant au niveau du coefficient “représentatif” du marché français (calculé avec le modèle développé par EIOPA (cf. section 2.10, équation 2.10), il est dépendant des derniers points à gros volume (en nombre non négligeable dans ce cas). Si nous suivons le formalisme de l’EIOPA consistant à ne pas tenir compte de ces points, le résultat serait de 8,7%. Si nous tenons compte de ces points, alors le résultat serait plutôt de 4,8%.

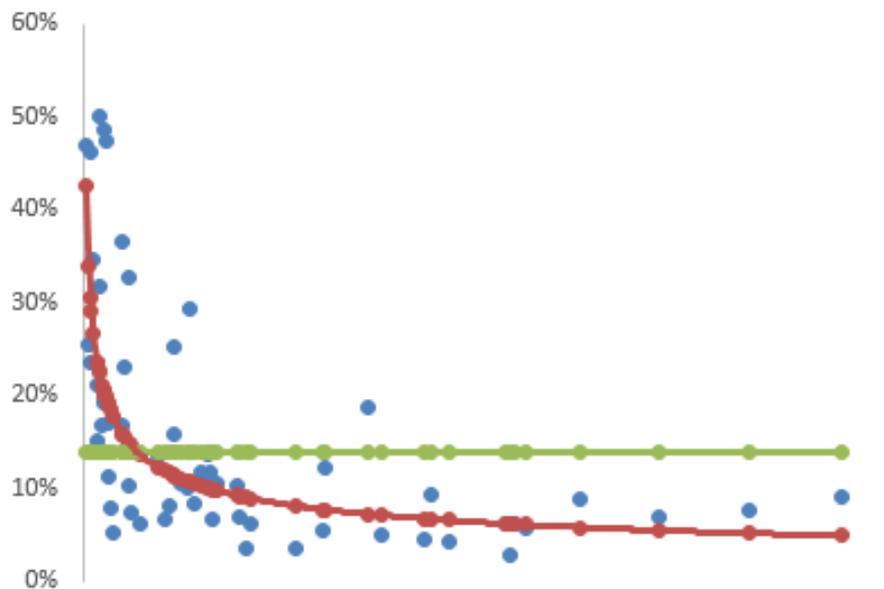


FIGURE 4.48 – Comparaison à la formule standard pour le segment “protection du revenu”

Le volume correspondant à l’intersection de la loi puissance (courbe rouge) avec le niveau de la formule standard (courbe verte) est d’environ 15M€. En séparant l’ensemble des sociétés du marché par rapport à ce volume et dans le cadre de ce modèle théorique, 90% du marché français devrait être désavantagé

par la formule standard et 10% devrait être avantagé (en termes de volume de BE de réserve).

Remarque 4.8.1 Pour une meilleure lisibilité du graphe, deux sociétés de volume important ont été exclues de l'échantillon, sans que cela ne change le résultat de la régression.

4.9 SEGMENT “ASSISTANCE”

Ce segment d'activité regroupe, comme indiqué dans la section 3.1, les engagements d'assurance couvrant l'assistance aux personnes qui rencontrent des difficultés lorsqu'elles sont en voyage, en déplacement ou loin de leur lieu de résidence habituel. Ce segment est à déroulement court.

Taille de l'échantillon

Le tableau ci-dessous détaille le nombre de sociétés utilisées lors de l'étude de ce segment après le retraitement de données (cf. section 3.3). Les 17 “triangles” qui ont été enlevés s'expliquent par une mauvaise qualité de données pour 9 cas et 6 autres à cause d'un volume trop faible.

Segment d'activité	“assistance”
Nr initial d'organismes	38
Nr organismes hors échantillon	17
Nr final	21
<i>Dont retraitements</i>	1

FIGURE 4.49 – Nombre de triangles analysés pour le segment d'activité “assistance”

Quelques indicateurs statistiques

Alors que l'ensemble des indicateurs statistiques sont homogènes, comme pour les autres lignes d'activités (cf. tableau 4.50), la décroissance en loi puissance du coefficient de variation par rapport au volume ressort de moins bonne qualité. La durée est courte, environ 1,25, comme pour le segment “autre assurance des véhicules à moteur” et les facteurs d'émergence sont très proches de 100% témoignant un risque qui se développe à l'ultime en une seule année. L'exposant de la loi puissance est presque nul (estimé à $-0,06$) et le coefficient de détermination est de $R^2 = 0,11$ (cf. figure 4.51). Pour rappel, un exposant nul est équivalent à une droite constante. La moyenne des coefficients de variation de l'échantillon est de 20%, ce qui est proche du calibrage de la formule standard 22%.

Le coefficient de détermination, qui traduit de la qualité de la régression linéaire, ressort très petit pour ce segment. Cela peut s'expliquer par le faible nombre de données (nous observons une forte sensibilité du coefficient en cas de rajout/suppression d'une données), et traduit également une forte hétérogénéité entre les données. Cette hétérogénéité entre les sociétés peut s'expliquer par la comptabilisation des frais, qui peut fortement différer d'un organisme à l'autre. Par exemple, il est possible que l'assureur soit amené à réaliser des

Statistique	Duration	Facteur d'émergence
Moyenne	1,24	98%
1er décile	1,08	93%
1er quartile	1,15	98%
3ème quartile	1,29	100%
9ème décile	1,41	100%

FIGURE 4.50 – Statistiques sur le segment "assistance"

dépenses au titre de l'activité assistance sans que finalement aucun sinistre ne soit survenu. Dans ce cas précis, comment l'assureur enregistre-t-il les frais? Les pratiques semblent varier d'un assureur à l'autre, entraînant ainsi une forte hétérogénéité au sein de ce segment.

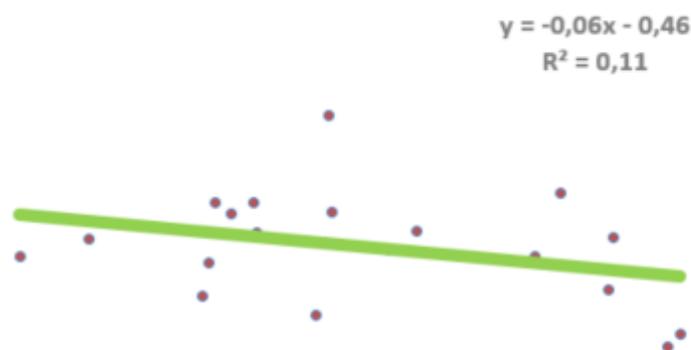


FIGURE 4.51 – Estimation du degré de la décroissance puissance

Finalement, aucune conclusion ne peut être levée pour le segment assistance, au vu du faible nombre de données et de la forte hétérogénéité qui en ressort.

Conclusions



L'objet principal de ce mémoire a été d'analyser des données du marché français de l'assurance afin d'étudier la nature de la décroissance du coefficient de variation (l'équivalent du « sigma » de la formule standard) pour le risque de réserve sur différents segments d'activité Solvabilité 2. Le modèle théorique simplifié où les risques sont indépendants et identiquement distribués revient à une décroissance en loi puissance avec un exposant de $-0,5$. La formule $3 \times \sigma \times V$ revient à considérer une fonction constante (puissance avec un exposant de 0).

Traitement des données utilisées

Les données utilisées ont été tirées des états quantitatifs Solvabilité 2, dits *Quantitative Report Templates* (ou QRT) qui sont remis au superviseur dans leur intégralité à fréquence annuelle et plus particulièrement les états s.19.01.01. Ces états contiennent, pour chaque société et chaque segment d'activité exercé, l'ensemble des “triangles” de liquidation. Nous avons ainsi utilisé les triangles de paiements cumulés et de charges, que ceux-ci soient bruts ou nets de réassurance. Notons que les sociétés publient dans le cadre du SFCR (*Solvency and Financial Conditions Reports*) un “triangle” global de liquidation des sinistres (qui est donc la somme de tous les segments d'activité).

La sélection des sociétés s'est déroulée en plusieurs étapes. Tout d'abord, nous avons sélectionné les sociétés¹ avec des provisions techniques non-vie supérieures à 35M€ (soit 99% du marché français en termes de provisions techniques non-vie). Les QRTs considérés pour ces sociétés sont ceux à fin 2019 (et à défaut 2018 ou 2017). En cas d'absence des états quantitatifs s.19.01.01 ou d'une taille de triangles inférieure à 4 ans, la société a finalement été exclue de l'étude. L'échantillon final représente presque 90% du marché français en termes de provisions techniques non-vie soit l'équivalent de 122G€ de meilleure estimation pour sinistres. Le tableau suivant présente le nombre de sociétés considérées par segment d'activité, une fois les retraitements décrits précédemment effectués :

Segment d'activité	Nr. d'organismes	BE de réserve
Frais médicaux	89	7,5G€
Protection du revenu	94	20G€
Indemnisation des travailleurs	3	0,3G€
Responsabilité civile automobile	70	30G€
Autres dommages auto	69	3,3G€
Assurance maritime, aérienne et transport	33	1,5G€
Dommages aux biens	76	19G€
Responsabilité civile générale	75	38G€
Crédit et cautionnement	4	0,1G€
Protection juridique	61	1G€
Assistance	38	0,4G€
Pertes pécuniaires diverses	46	1G€

FIGURE 5.1 – Nombre d'organismes par segment d'activité

À noter que les sociétés de réassurance n'ont pas été incluses pour deux

1. Nous avons utilisé le mot “société” pour tout les types d'entreprise d'assurance : sociétés anonymes, instituts de prévoyance ou mutuelles

raisons. Tout d'abord, les lignes d'activité de réassurance non proportionnelle n'ont pas été étudiées du fait du faible nombre de sociétés et de la concentration de ce type d'activité. En ce qui concerne l'activité de réassurance proportionnelle, les risques ne sont pas limités au marché français et ressortent donc hétérogènes. Une étude approfondie pour les réassureurs aurait été souhaitable, mais nous ne disposons pas de données suffisantes pour une telle étude.

Trois segments d'activité ont été exclus du cadre de l'étude : l'*indemnisation des travailleurs* car ce segment n'est pas pratiqué en France du fait du recours à la sécurité sociale, le *crédit et cautionnement* à cause du faible nombre de données et le segment *assurance maritime, aérienne et transport* à cause de la forte concentration (trois sociétés représentent environ 85% du volume).

L'échantillon total de l'étude comporte 618 triangles. L'étude individuelle de chaque triangle aurait été très coûteuse en temps. Nous avons ainsi développé un outil qui prend en entrée les triangles de liquidation et construit des indicateurs statistiques afin d'identifier les problèmes de qualité de données. En cas de problème identifié, les triangles ont été analysés individuellement, 77 cas parmi les 618. Dans 29 cas parmi ces 77 nous avons retraité les données (suppression de certaines parties du triangle) mais seulement pour les raisons suivantes : rupture dans le provisionnement, année atypique, acquisition ou vente de portefeuille, forte croissance ou baisse du chiffre d'affaires ou erreurs manifestes de données. Ces raisons peuvent perturber un triangle de manière à ce que la volatilité estimée ne reflète pas le vrai profil de risque de la société. Afin de définir l'échantillon final, trois filtres ont été utilisés : volume faible ne garantissant pas l'application de la loi des grands nombres, profondeur d'historique insuffisante pour des risques à développement long comme ceux de responsabilité civile, temps de règlement atypique afin d'exclure les garanties d'assurance construction. Au final, l'étude a porté sur 440 triangles, soit une moyenne 49 par segment d'activité.

Décroissance empirique du risque par rapport au volume

Notons également que les triangles utilisés *in fine* sont ceux des paiements bruts et nets de réassurance. Les triangles de charges posent des problèmes de qualité des données que nous expliquons par le niveau de prudence du provisionnement dossier-dossier globalement significatif, variable dans le temps et différent d'une société à l'autre. De plus, à des fins de comparaison, nous avons fait un choix similaire à celui de l'EIOPA lors du calibrage de la formule standard. Les résultats entre les triangles de paiements bruts et nets sont très proches pour l'ensemble des segments d'activité étudiés et nous indiquerons ceux relatifs aux paiements bruts. Pour la mesure de volume, nous avons eu le choix entre la meilleure estimation des sinistres et la "provision *Chain-Ladder*" (i.e. la provision obtenue par l'application directe et sans aucun retraitement de la méthode *Chain-Ladder*). Nous avons opté pour la seconde option pour deux raisons. La meilleure estimation des sinistres contient également des frais de gestion de sinistres et d'administration qui ne sont pas identiques entre les sociétés et cela aurait pu introduire un biais. De plus, la "provision *Chain-Ladder*" est la moyenne implicite des modèles mathématiques que nous avons utilisés

pour l'estimation du risque.

Le calcul de l'écart-type lié au risque de réserve a été calculé en considérant plusieurs méthodes : le modèle de Mack, le modèle de Merz & Wüthrich, ou encore par la technique de *bootstrap*. La méthode privilégiée a été le modèle de Merz & Wüthrich car il estime un risque à un an, ce qui est la mesure du risque du régime prudentiel actuel. Quant au modèle de Mack, il estime un risque dit à l'ultime. La technique de *bootstrap* implémentée pour estimer un risque à un an donne des résultats proches de la méthode de Merz & Wüthrich sur des données statistiques de bonne qualité (i.e. sur l'échantillon final).

Sur l'ensemble des segments d'activité nous observons une décroissance du coefficient de variation en fonction du volume. Le tableau ci-dessous indique pour chaque ligne d'activité l'exposant de la décroissance en loi puissance ainsi que les coefficients de détermination relatifs à la régression utilisée pour l'estimation de l'exposant. L'exposant de la décroissance en loi puissance se trouve entre le modèle théorique simplifié et l'hypothèse de la formule standard (i.e. entre $-0,5$ et 0). D'autres types de décroissance ont été testés, logarithmique et exponentielle, mais les coefficients de détermination relatifs aux régressions sont de moins bonne qualité (i.e. plus proches de 0).

Segment d'activité	Exposant	R2
Frais médicaux	-0,18	0,24
Protection du revenu	-0,40	0,56
Responsabilité civile automobile	-0,36	0,76
Autres dommages auto	-0,19	0,39
Dommages aux biens	-0,27	0,46
Responsabilité civile générale	-0,22	0,47
Protection juridique	-0,29	0,56
Assistance	-0,06	0,11
Pertes pécuniaires diverses	-0,21	0,45

FIGURE 5.2 – L'exposant de la décroissance du coefficient de variation par rapport au volume par segment d'activité

Le segment de responsabilité civile automobile est celui où la qualité de la régression est la meilleure et nous l'expliquons par l'homogénéité des risques. En effet, ce segment d'activité contient une seule garantie, elle est obligatoire et donc encadrée par le Code des Assurances. Cela rend la ligne d'activité homogène au sein d'une société mais aussi entre les sociétés du marché français. Nous observons que les segments pour lesquels les exposants sont les plus proches du modèle théorique simplifié sont ceux de responsabilité civile automobile et de protection de revenu. La qualité de la régression est de moins bonne qualité sur le segment de protection de revenu car deux types de risques sont mélangés, l'assurance emprunteur et les garanties d'incapacité-invalidité.

Le segment d'assistance présente un exposant atypique, proche de 0 et c'est le segment avec la plus forte hétérogénéité des résultats. Nous l'expliquons par le faible nombre de données mais aussi par la comptabilisation des frais qui peut fortement différer d'un organisme à l'autre. Les assureurs sont amenés à réaliser des dépenses au titre de l'activité assistance sans que finalement aucun sinistre ne soit survenu. Pour le reste des segments d'activité, les exposants sont

d'environ $-0,2$ avec une qualité de la régression raisonnable autour de 50%. Le segment d'activité de frais médicaux ressort avec une qualité de la régression assez faible. Les coefficients de variation sont assez hétérogènes sur le marché français et nous pensons que c'est un problème de qualité de données.

Risque à un an et risque à l'ultime

Des exposants avec une décroissance en loi puissance similaire sont obtenus en utilisant une estimation du risque à l'ultime avec le modèle de Mack. En effet, le facteur d'émergence, défini comme le rapport entre le risque à un an (obtenu via la méthode de méthode de Merz & Wüthrich) et le risque à l'ultime (obtenu via la méthode Mack) est un indicateur stable sur l'ensemble du marché au sein de chaque segment de risque. Ce facteur d'émergence permet de comprendre quelle partie du risque à l'ultime émerge pendant la première année de la liquidation des réserves. A titre d'exemple, pour les segments avec une liquidation rapide (tels que les frais médicaux ou les dommages auto), le risque à l'ultime est très proche du risque à un an et cela se traduit par des facteurs d'émergence très proches de 100%. L'étude menée montre que de manière générale le facteur d'émergence est anti-corrélé avec la durée (i.e. le temps moyen de liquidation des réserves). En effet, plus la durée est grande, plus le facteur d'émergence est en général faible.

Segment d'activité	Duration	Facteur d'émergence
Frais médicaux	1,20	97%
Protection du revenu	2,50	85%
Resp. civile auto	3,27	67%
Autres dommages auto	1,27	94%
Dommages aux biens	1,69	86%
Resp. civile générale	3,79	71%
Resp. civile générale	3,79	71%
Protection juridique	2,46	85%
Assistance	1,24	98%
Pertes pécuniaires diverses	1,76	92%

FIGURE 5.3 – Comparaison entre la formule standard et le coefficient "représentatif" du marché français par segment d'activité pour le risque de réserve

Comparaison avec la formule standard

Pour le calcul des coefficients de variation par segment d'activité (i.e. les chocs de la formule standard), EIOPA a utilisé un modèle de minimisation de l'erreur quadratique :

$$\sum_c \left(\sigma_{lob} V_{c,lob} - \sqrt{MSEP(R)_{c,lob}} \right)^2, \text{ où :}$$

- c est un indice qui prend ses valeurs dans l'ensemble des sociétés de l'échantillon ;
- $V_{c,lob}$ est une mesure de volume et qui a été considérée égale au résultat de l'application de la méthode de Chain-Ladder ;
- $MSEP(R)_{c,lob}$ est une mesure de l'erreur quadratique (en fonction du modèle utilisé).

En pratique, en résolvant un problème de minimum absolu, le problème de minimisation ci-dessus devient :

$$\hat{\sigma}_{lob} = \frac{\sum_c V_{c,lob} \sqrt{MSEP(R)_{c,lob}}}{\sum_c V_{c,lob}^2} = \frac{\sum_c V_{c,lob}^2 CoV_{c,lob}}{\sum_c V_{c,lob}^2}.$$

Cette formule revient à faire une moyenne pondérée des coefficients de variation individuels par le carré des volumes. Le calibrage est donc fortement dépendant des gros volumes et EIOPA les a enlevés systématiquement lors du calibrage de la formule standard. Le tableau ci-dessous indique les coefficients représentatifs du marché français (calculé avec la formule développée par EIOPA). Nous indiquons à chaque fois deux valeurs : une en gardant l'ensemble de l'échantillon et une en suivant le formalisme de l'EIOPA qui consiste à enlever les gros volumes. Si les deux sont identiques, cela signifie une absence des gros volumes. Avant de commenter le tableau, notons tout d'abord que le choix final des coefficients de variation retenus par EIOPA est le résultat d'une concertation européenne incluant l'ensemble des superviseurs nationaux mais également les assureurs.

Segment d'activité	FS	CoV tout marché	CoV sauf gros volumes
Frais médicaux	5,7%	6,7%	6,7%
Protection du revenu	14%	4,8%	8,7%
Resp. civile auto	9%	4,4%	8,8%
Autres dommages auto	8%	7,3%	11,2%
Dommages aux biens	10%	8,6%	12,4%
Resp. civile générale	11%	11,2%	11,2%
Protection juridique	5,5%	5%	5,8%
Assistance	22%	20%	20%
Pertes pécuniaires diverses	20%	20%	20%

FIGURE 5.4 – Comparaison entre la formule standard et le coefficient "représentatif" du marché français par segment d'activité pour le risque de réserve

Ce tableau montre d'abord l'effet important des gros volumes pour les segments d'activité qui sont concernés. De manière générale la formule standard semble adaptée pour le marché français de l'assurance. Les seules exceptions sont le segment des frais médicaux où la formule standard donne un choc plus faible alors que pour le segment de protection de revenu c'est l'inverse. Pour le reste des segments d'activité nous observons le principe de Pareto : environ 75% à 80% du marché en termes de provisions prudentielles est plutôt désavantagé par la formule standard et 20% à 25% du marché est avantagé par la formule standard alors qu'en termes de nombre d'entreprises c'est l'inverse.

Comparaison avec la méthode réglementaire pour le calcul des paramètres propres

En ce qui concerne le calcul des paramètres propres pour remplacer ceux de la formule standard (couramment appelés USP pour *Undertaking specific parameters*), deux méthodes statistiques relatives à leur évaluation sont stipulées à l'annexe XVII du règlement délégué. Parmi les deux, une est très similaire à celle de Merz & Wüthrich, une approximation étant faite, à la baisse, pour simplifier la formule originale. Cette approximation, qui n'est rien d'autre qu'un développement limité de premier ordre, a des conséquences significatives, d'environ 15%,

pour l'estimation du risque pour les segments d'activité de responsabilité civile. Le tableau suivant montre la différence moyenne entre la méthode de Merz & Wüthrich et sa variante simplifiée (méthode USP). Les écarts sont proches au sein d'une même ligne d'activité sur l'ensemble du marché.

Segment d'activité	Écart relatif
Frais médicaux	1%
Protection du revenu	6%
Resp. civile auto	14%
Autres dommages auto	2%
Dommmages aux biens	5%
Resp. civile générale	14%
Protection juridique	6%
Assistance	1%
Pertes pécuniaires diverses	4%

FIGURE 5.5 – *Écart relatif entre la méthode de Merz & Wüthrich et sa variante simplifiée (méthode USP) par segment d'activité*



Bibliographie

- [1] R. Efron, B. et Tibshiran. *An Introduction to the Bootstrap*. Number 57 in Monographs on Statistics and Applied Probability. Chapman & Hall/CRC, 1993. (Cité page 14.)
- [2] EIOPA. Calibration of the premium and reserve risk factors in the standard formula of solvency ii. 2011. (Cité pages 24 et 27.)
- [3] EIOPA. Code des assurances. 2020. (Cité pages 6 et 22.)
- [4] R. England, P. et Verral. A flexible framework for stochastic claims reserving. 2001. (Cité page 14.)
- [5] R. England, P. et Verral. Stochastic claims reserving in general insurance. 8 :443–518, 2002. (Cité page 9.)
- [6] Parlement Européen. Directive 2009/138/ce du parlement européen et du conseil. 2009. (Cité pages 6, 18 et 20.)
- [7] Commission Européenne. Règlement délégué (ue) 2015/35 de la commission européenne. 2014. (Cité pages 6, 7, 18, 19 et 22.)
- [8] Commission Européenne. Règlement délégué (ue) 2019/981 de la commission européenne. 2019. (Cité page 20.)
- [9] T. Mack. Distribution-free calculation of the standard error of chain ladder reserve estimates. *Astin Bulletin*, 23, 1993. (Cité pages 9 et 11.)
- [10] Merz M. Wüthrich, M. Modelling the claims development result for solvency purposes. cas e-forum (fall. 2008. (Cité page 23.)
- [11] Merz M. et Lysenko N. Wüthrich, M. Uncertainty of the claims development result in the chain ladder method. *Scandinavian Actuarial Journal*, 1 :63–84, 2009. (Cité page 12.)