

MODÈLES D'EXTRAPOLATION DE LA MORTALITÉ AUX GRANDS ÂGES

AKI QUASHIE & MICHEL DENUIT

Institut des Sciences Actuarielles & Institut de Statistique
Université Catholique de Louvain
Louvain-la-Neuve, Belgium

13 février 2005

Résumé

Cet article compare différentes méthodes de fermeture des tables de mortalité sur base de données relatives aux populations française, néerlandaise et suédoise. Les mérites respectifs de ces méthodes sont soigneusement comparés. Les conséquences de leur utilisation en matière de tarification de rentes viagères sont examinées en détail.

Mots-clé: assurance vie, tables de mortalité, âge ultime, fermeture, rentes viagères.

1 Introduction

En 2001, 72 millions des 6.1 milliards d'habitants de la planète avaient plus de 80 ans selon les Nations Unies. Les "80 ans et plus" constituaient alors 1.2% de la population mondiale. Alors qu'on s'attend à une croissance d'environ 50% de la population mondiale d'ici 2050 (pour atteindre 9.3 milliards d'habitants), le nombre de personnes âgées de 80 ans et plus atteindrait d'ici là 379 millions. Dans les pays industrialisés, ce nombre va vraisemblablement tripler, passant de 38 millions en 2001 à 113 millions en 2050. Alors que de plus en plus de nos contemporains atteignent des âges très élevés, la mortalité à ces âges est cependant encore mal connue.

Des études scientifiques récentes ont mis en lumière l'allure remarquable de la mortalité aux grands âges, à savoir une décroissance du taux de croissance de la mortalité à partir d'un certain âge. HORIUCHI & WILMOTH (1998) ont ainsi montré que sur l'échelle logarithmique, la courbe des taux de mortalité présentait une allure concave aux grands âges. Ceci a conduit les actuaires et les démographes à rechercher des modèles compatibles avec cette réalité. De nombreux modèles ont été développés. Voyez par exemple BOLESZAWSKI & TABEAU (2001), THATCHER (1999) et THATCHER ET AL. (1998). Nous avons retenu pour cette étude le modèle Logistique et une de ses variantes, le modèle de Kannistö, le modèle de LINDBERGSON (2001), le modèle de DENUIT & GODERNIAUX (2005), ainsi que les modèles très largement utilisés de COALE & GUO (1989) et COALE & KISKER (1990). L'objet de cet article est double

- d'une part comparer les méthodes d'extrapolation de la mortalité aux grands âges mentionnées plus haut sur base de jeux de données; et
- d'autre part mesurer l'impact de ces modélisations sur la tarification des rentes viagères.

Dans ce travail, nous retenons l'hypothèse de constance par morceaux des taux instantanés de mortalité, i.e.

$$\mu_{x+\xi} = \mu_x \text{ pour } 0 \leq \xi < 1 \text{ et } x \text{ entier .} \quad (1.1)$$

En notant comme de coutume q_x la probabilité de décès à l'âge x et $p_x = 1 - q_x$ la probabilité de survie correspondante, on vérifie facilement que dans ce cas

$$\mu_x = -\ln p_x \quad (1.2)$$

Les données utilisées se rapportent aux 3 pays suivants: Suède, Pays-Bas et France respectivement pour les années 2002, 2001 et 1999. Les données sont prises dans la Base de Données de la Mortalité Humaine (HMD pour "Human Mortality Database"). Le lecteur intéressé pourra avantageusement se rendre sur le site internet *www.mortality.org* pour avoir de plus amples détails sur la description de la base de données, l'historique du projet, l'explication des données, les équipes de recherches impliquées, les données relatives à chaque pays, les différents formats de données, etc.

La Figure 1.1 représente les q_x bruts sur l'échelle logarithmique pour l'ensemble des données. On retrouve une allure désormais classique caractérisée par les phénomènes suivants:

- la forte mortalité post-natale;
- le creux de la mortalité des 5-15 ans;

- la bosse de la mortalité (encore appelée "bosse accidents") vers 20-30 ans. Elle est due aux morts violentes (accidents et suicides) dans cette tranche d'âges, et est nettement plus marquée chez les hommes que chez les femmes;
- la mortalité inférieure des femmes à partir de 15 ans environ;
- l'allure générale quasi linéaire après 40 ans;
- la légère allure concave en fin de table.

La suite de cet article est organisée de la manière suivante. La section 2 présente chacune des méthodes d'extrapolation de la mortalité aux grands âges, et donne les résultats de leur application sur l'ensemble des données de la Figure 1.1. La section 3 compare les méthodes entre elles et examine les conséquences de leur utilisation en matière de tarification des rentes viagères.

2 Méthodes d'extrapolation de la mortalité aux grands âges

2.1 La méthode de Denuit & Goderniaux

La méthode de DENUIT & GODERNIAUX (2005) travaille sur les quotients de mortalité et introduit une contrainte de fermeture des tables de mortalité. Elle consiste à ajuster par moindres carrés le modèle log-quadratique

$$\ln \widehat{q}_x = a + bx + cx^2 + \epsilon_x \text{ avec } \epsilon_x \text{ iid } \mathcal{N}or(0, \sigma^2) \quad (2.1)$$

sur base des observations relatives aux âges les plus élevés, avec les deux contraintes

$$q_{130} = 1 \quad (2.2)$$

$$q'_{130} = 0. \quad (2.3)$$

où q'_x désigne la dérivée première de q_x considéré comme fonction de l'âge x . Les contraintes (2.2) et (2.3) conduisent à

$$a + bx + cx^2 = c(130^2 - 260x + x^2) \quad (2.4)$$

Les contraintes (2.2) et (2.3) imposent à la courbe des quotients de mortalité une allure concave aux grands âges et l'existence d'une tangente horizontale au point $x = 130$ ans. La contrainte (2.3) a pour conséquence d'empêcher une décroissance éventuelle de q_x aux âges avancés.

Afin de déterminer l'âge à partir duquel les \widehat{q}_x bruts seront remplacés par les valeurs ajustées provenant du modèle (2.1), nous avons choisi comme critère la maximisation du coefficient de détermination R^2 (en faisant varier cet "âge de raccord" dans l'intervalle 50-85 ans). Un lissage est parfois nécessaire autour de l'"âge de raccord". Une simple moyenne géométrique pour les âges $x = x_0 - 5, \dots, x_0 + 5$ où x_0 est l'"âge de raccord" suffit en général. Ceci évite une cassure au voisinage de cet âge. La Figure 2.1 reprend les quotients de mortalité bruts et prolongés pour la tranche d'âges 80-120 ans. Le modèle (2.1)-(2.4) semble prolonger naturellement les quotients de mortalité aux âges élevés.

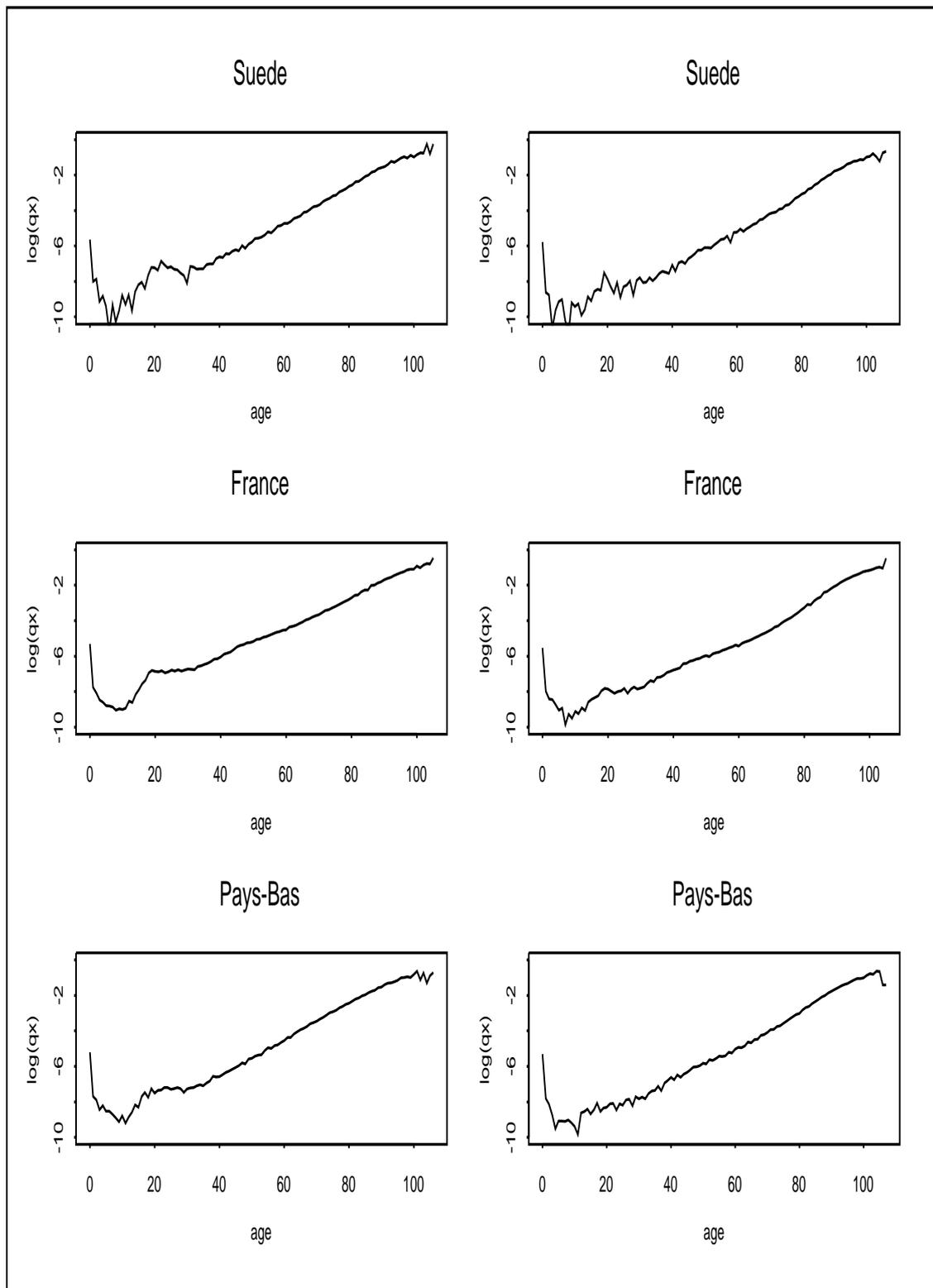


FIG. 1.1 – Probabilités de décès brutes \hat{q}_x en fonction de l'âge x sur l'échelle logarithmique pour la Suède, la France et les Pays-Bas (les hommes sont à gauche, les femmes à droite)

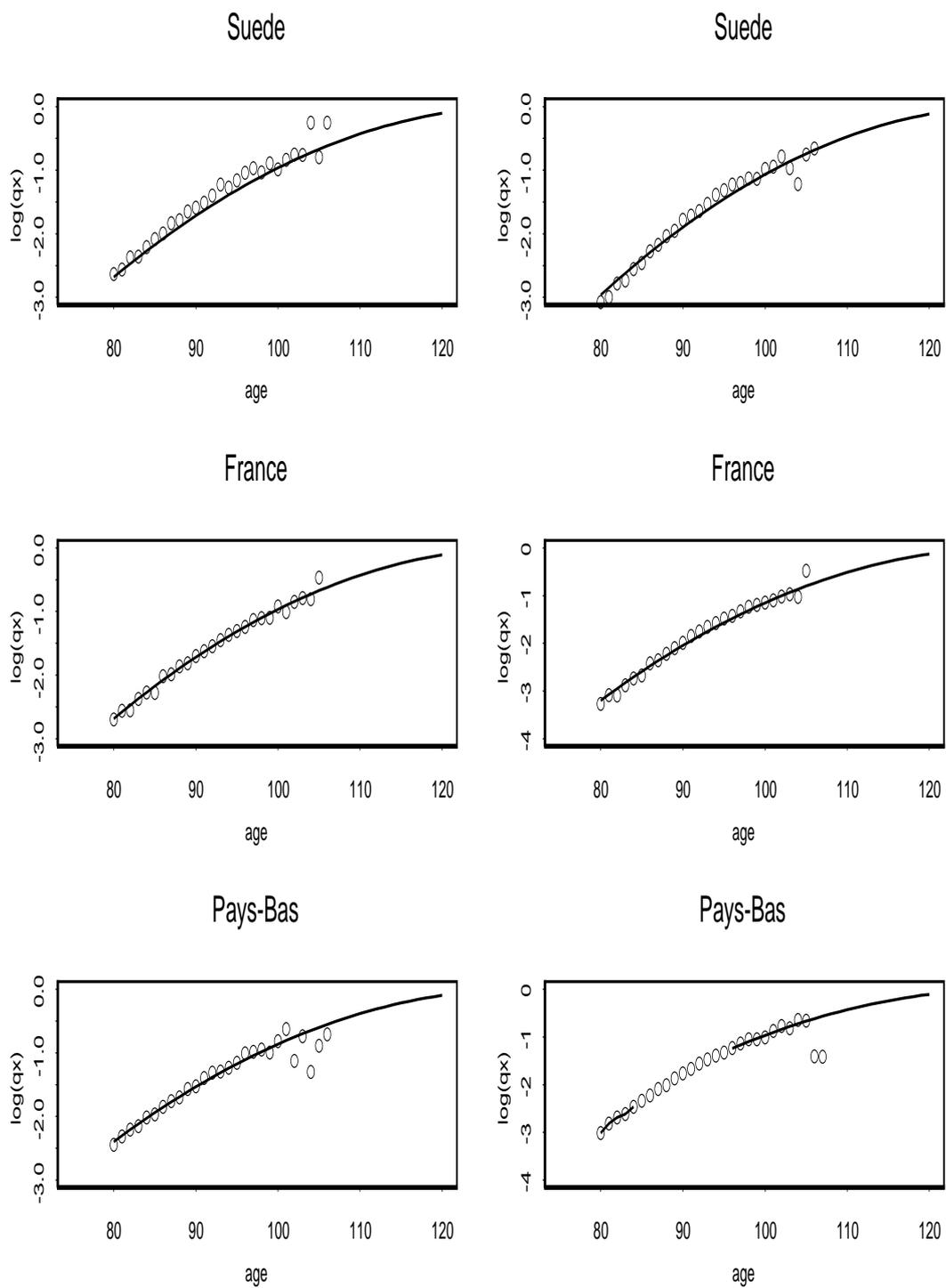


FIG. 2.1 – Comparaison des quotients de mortalité bruts et des quotients de mortalité reconstitués et extrapolés par la méthode de Denuit & Goderniaux

2.2 La méthode de Coale & Guo

COALE & GUO (1989) ont proposé une technique de fermeture des tables de mortalité dans laquelle les taux de mortalité observés aux grands âges (85 ans et plus) sont remplacés par une séquence de taux de mortalité extrapolés aux groupes d'âges 85-89, 90-94, ..., 105-109 ans. Sur base de données relatives à 6 pays (Autriche, France, Japon, Norvège, Pays-Bas, Suède), les auteurs ont constaté une décroissance constante du taux de croissance de la mortalité d'un intervalle d'âges de 5 ans à un autre après 80 ans ou 85 ans. Cette constatation est à la base de la méthode.

Les auteurs ont travaillé avec des données par intervalles d'âges de 5 ans. Nous nous proposons ici de donner une version adaptée du modèle de COALE & GUO (1989) qui donne les taux de mortalité par âge. En posant $k' = \ln\left(\frac{\hat{\mu}_{75}}{\hat{\mu}_{74}}\right)$, ces auteurs extrapolent les taux de mortalité aux âges élevés à l'aide de la formule

$$\hat{\mu}_{75+i} = \hat{\mu}_{75} \exp\left(ik' - \frac{i(i+1)}{2}R\right), \quad i = 1, \dots, 34. \quad (2.5)$$

En tenant compte de la contrainte

$${}_5\hat{\mu}_{105} - {}_5\hat{\mu}_{75} = 0.66, \text{ où } {}_5\hat{\mu}_x \approx \prod_{i=0}^4 (\hat{\mu}_{x+i})^{\frac{1}{5}}, \quad (2.6)$$

imposée arbitrairement par les auteurs, il vient finalement

$$R = \frac{30k' - \ln\left(\frac{0.66 + {}_5\hat{\mu}_{75}}{{}_5\hat{\mu}_{75}}\right)}{525}.$$

La Figure 2.2 compare les taux bruts $\hat{\mu}_x$ et les valeurs obtenues grâce à l'équation (2.5) sur l'intervalle d'âges 80-120 ans. On constate que les courbes ajustées s'éloignent parfois considérablement des valeurs observées.

2.3 La méthode de Coale & Kisker

Le modèle de COALE & KISKER (1990) postule que

$$\hat{\mu}_x = \hat{\mu}_{65} \exp\left(k_x \cdot (x - 65)\right) \quad \text{pour } x \geq 65 \quad \text{ans}, \quad (2.7)$$

où k_x est le taux de croissance de la mortalité à l'âge x . Les auteurs, après avoir examiné différents graphiques relatifs au taux de croissance de la mortalité k_x de 7 pays considérés comme ayant des données de mortalité fiables aux âges élevés (Autriche, ex-Allemagne de l'ouest, France, Japon, Pays-Bas, Norvège, Suède), ont remarqué que les k_x possédaient en général un pic aux alentours de 80 ans avant de décroître linéairement. Ceci les a amené à supposer que

$$k_x = k_{80} + s \cdot (x - 80), \quad x \geq 80. \quad (2.8)$$

Afin de définir le coefficient s relatif à la pente, COALE & KISKER (1990) fixent arbitrairement le taux de mortalité pour un âge ultime, soit 110 ans:

$$\hat{\mu}_{110} = \begin{cases} 1.0 & \text{pour les hommes} \\ 0.8 & \text{pour les femmes.} \end{cases} \quad (2.9)$$

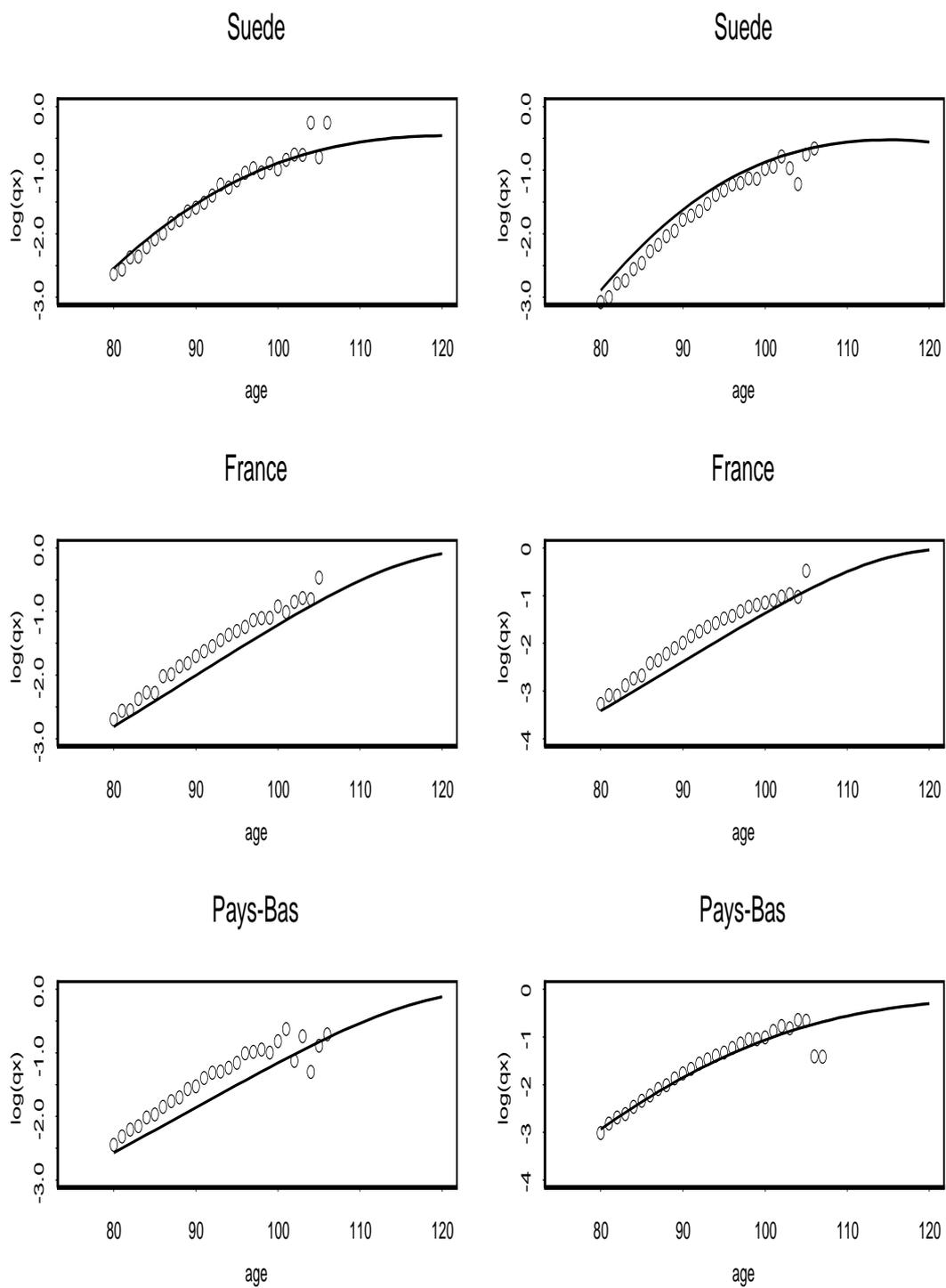


FIG. 2.2 – Comparaison des quotients de mortalité bruts et des quotients de mortalité reconstitués et extrapolés par la méthode de Coale & Guo

On a alors

$$\hat{\mu}_{110} = \hat{\mu}_{79} \exp \left(\sum_{x=80}^{110} k_x \right) = \hat{\mu}_{79} \exp \left(\sum_{x=80}^{110} (k_{80} + s.(x - 80)) \right)$$

qui donne

$$s = -\frac{\ln(\frac{\hat{\mu}_{79}}{\hat{\mu}_{110}}) + 31k_{80}}{465}.$$

Les taux de mortalité au-delà de 80 ans sont alors calculés de la manière suivante:

$$\hat{\mu}_x = \hat{\mu}_{79} \exp \left(\sum_{y=80}^x (k_{80} + s.(y - 80)) \right), \quad x \in \{80, 81, \dots, 109\}$$

ou encore

$$\hat{\mu}_x = \hat{\mu}_{x-1} \exp(k_{80} + s.(x - 80)), \quad x \in \{80, 81, \dots, 109\}.$$

Notez que les taux au-delà de 80 ans sont déterminées uniquement par $\hat{\mu}_{79}$ et $k_{80} = \frac{\ln(\frac{\hat{\mu}_{80}}{\hat{\mu}_{65}})}{15}$.

La Figure 2.3 révèle que les q_x bruts sont inférieurs aux q_x calculés par la méthode de COALE & KISKER (1990). Afin d'apprécier l'impact de la contrainte de fermeture (2.9), nous avons cherché à voir dans quelle mesure une modification de la valeur de $\hat{\mu}_{110}$ change les résultats. La Figure 2.4 obtenue pour les données suédoises illustre cela. On voit clairement que le choix de la valeur de $\hat{\mu}_{110}$ a un impact significatif, les courbes les plus proches des courbes brutes étant celles correspondant à $\hat{\mu}_{110} = 1.3$ pour les hommes et $\hat{\mu}_{110} = 1.1$ pour les femmes.

2.4 La méthode de Lindbergson

En travaillant sur des données suédoises relatives aux périodes 1988-92 et 1993-97, LINDBERGSON (2001) a proposé une technique qui consiste à modifier la formule de Makeham en remplaçant la croissance exponentielle par une droite aux âges élevés. Plus précisément, le taux à l'âge x est supposé de la forme

$$\mu_x(a, b, c, k) = \begin{cases} a + b. \exp(cx) & \text{si } x \leq \omega \\ a + b. \exp(c\omega) + k(x - \omega) & \text{si } x > \omega \end{cases} \quad (2.10)$$

où les paramètres a , b , c et k sont inconnus et ω est l'âge charnière où la mortalité tend à se stabiliser. Sous l'hypothèse de constance par morceaux du taux instantané de mortalité, on fixe alors ω sur base des observations et les paramètres a , b , c et ω intervenant dans (2.10) minimisent

$$Q^2 = \sum_x \frac{(\hat{\mu}_x - \mu_x(a, b, c, k))^2}{\hat{L}_x}$$

où $\hat{\mu}_x$ et \hat{L}_x sont respectivement, à l'âge x , le taux instantané de mortalité et l'exposition totale au risque observés. On répète alors la minimisation pour différentes valeurs de ω pour trouver la valeur ω_0 de ω qui minimise Q^2 . Précisons que les calculs sont effectués à partir d'un âge que nous noterons x_0 où les taux de mortalité sont croissants.

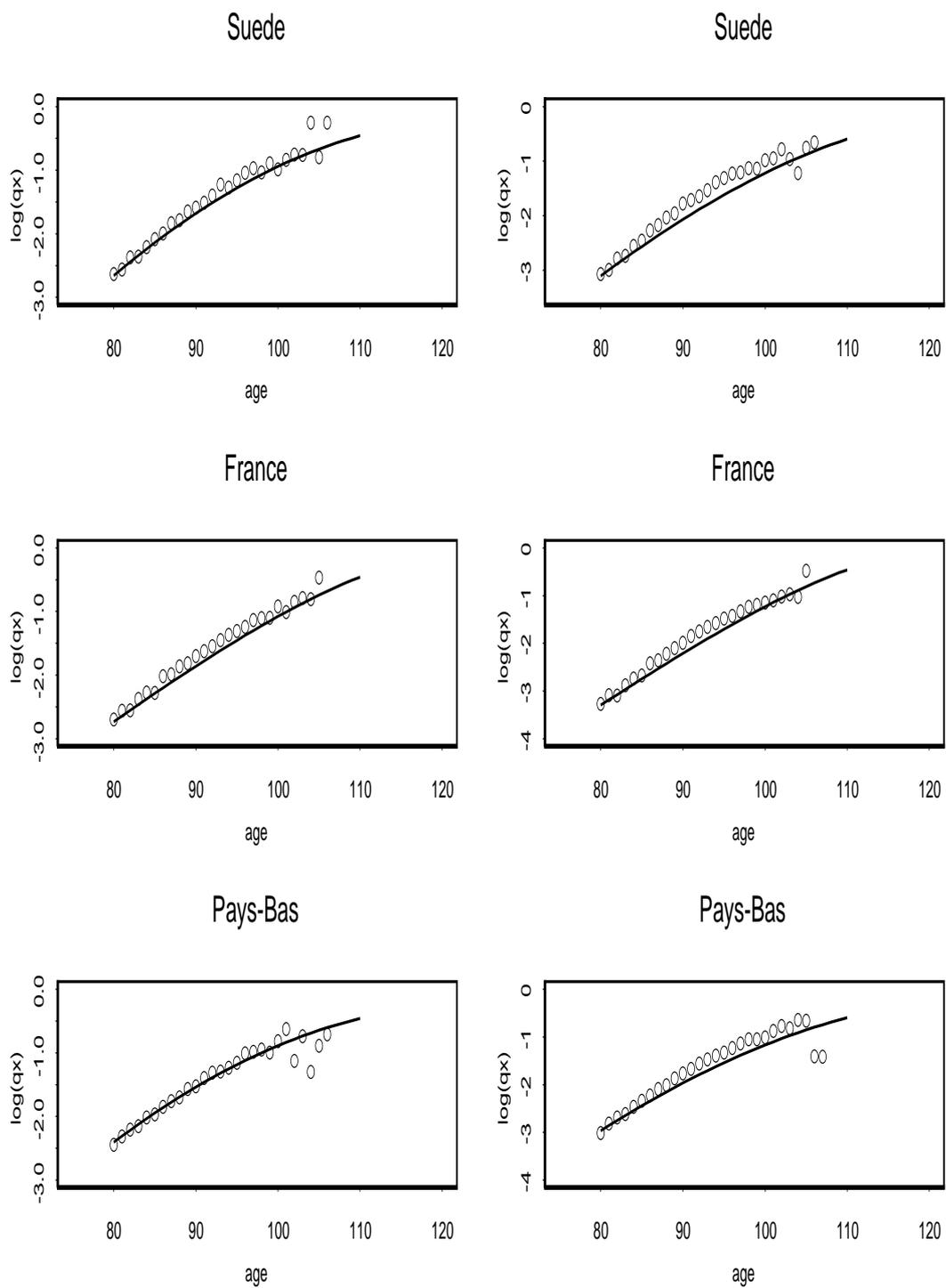
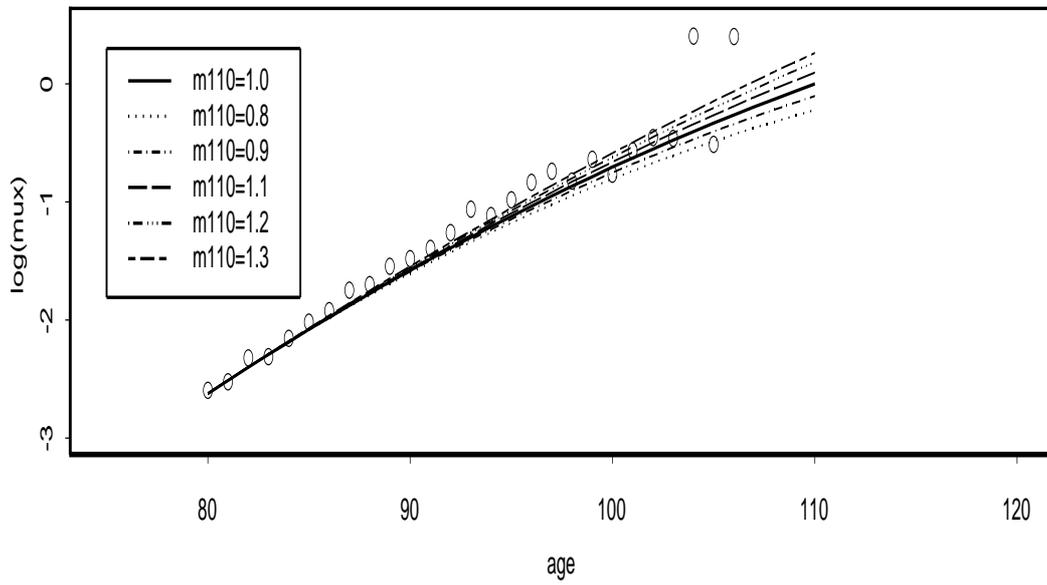


FIG. 2.3 – Comparaison des quotients de mortalité bruts et des quotients de mortalité reconstitués et extrapolés par la méthode de Coale & Kisker

Hommes



Femmes

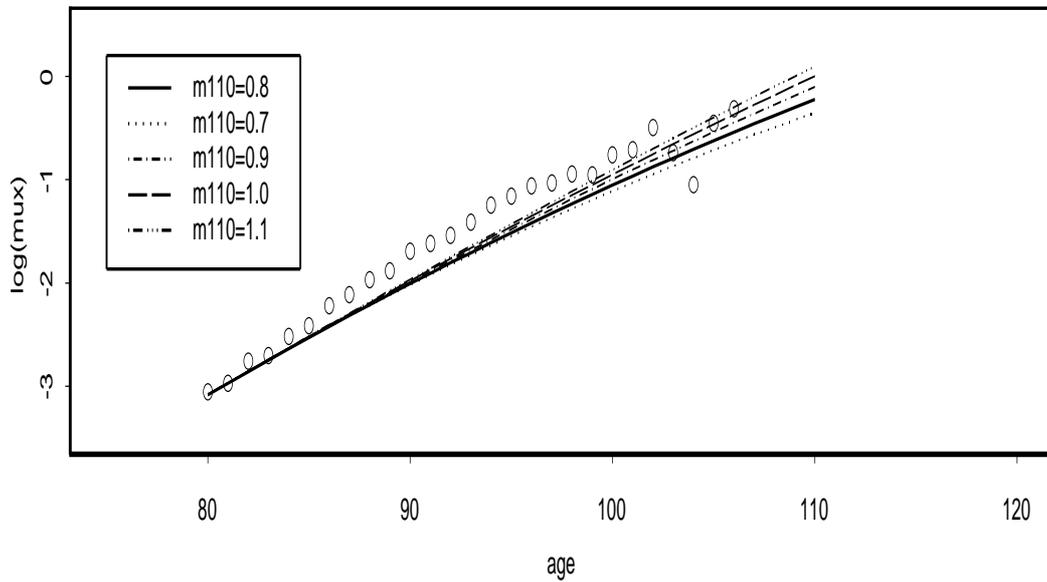


FIG. 2.4 – Taux de mortalité aux grands âges par la méthode de Coale & Kisker avec $\hat{\mu}_{110} \in \{0.8, 0.9, 1.0, 1.1, 1.2, 1.3\}$ pour les hommes suédois et $\hat{\mu}_{110} \in \{0.7, 0.8, 0.9, 1.0, 1.1\}$ pour les femmes suédoises

La Figure 2.5 représente sur l'échelle logarithmique les q_x bruts et ajustés par la technique de LINDBERGSON (2001) pour l'ensemble des données. On peut remarquer une allure rectiligne aux âges élevés qui contraste avec une allure concave classique et s'éloigne parfois sensiblement des données observées. Il semble donc que modéliser les taux aux âges élevés par une fonction affine ne soit pas judicieux.

2.5 Le modèle Logistique

La courbe d'une fonction logistique est une courbe "en S", c'est-à-dire qui croît d'abord de manière accélérée puis de manière ralentie. De plus, elle a un comportement asymptotique intéressant. Le modèle logistique postule que

$$\mu_x = c + \frac{ae^{bx}}{1 + \sigma^2 \frac{a}{b}(e^{bx} - 1)} \quad (2.11)$$

Nous avons ajusté le modèle aux données par la moindres carrés dans la tranche d'âges 85-98 ans. Ceci a donné les résultats présentés à la Figure 2.6. On constate une bonne adhérence aux données brutes au-delà de 99 ans.

2.6 Le modèle de Kannistö

Ce modèle est un cas particulier du modèle logistique stipulant que

$$\hat{\mu}_x = \frac{ae^{bx}}{1 + a(e^{bx} - 1)}. \quad (2.12)$$

Par rapport à (2.11), on a $c = 0$ et $\sigma^2 = b$.

Ce modèle a l'avantage de n'avoir que 2 paramètres à estimer (a et b), ce qui le rend facile à ajuster. Dans ce modèle, les taux sont limités à la valeur 1 tant pour les hommes que les femmes.

L'application du modèle de Kannistö à l'ensemble des données donne les résultats présentés à la Figure 2.7. On constate une bonne adhérence aux données brutes au-delà de 99 ans.

3 Comparaison des méthodes

Les mauvaises performances des modèles de Coale & Kisker et de Coale & Guo (constatées aux Figures 2.2-2.3) nous amènent à ne retenir pour la suite de notre travail que la méthode de DENUIT & GODERNIAUX, le modèle Logistique, le modèle KANNISTÖ et la méthode de LINDBERGSON.

3.1 Comparaison sur base des q_x

La Figure 3.1 représente dans l'intervalle d'âges 80-120 ans (sur l'échelle logarithmique) les q_x bruts et les q_x obtenus par chacune des 4 techniques que nous avons retenues (où "DG",

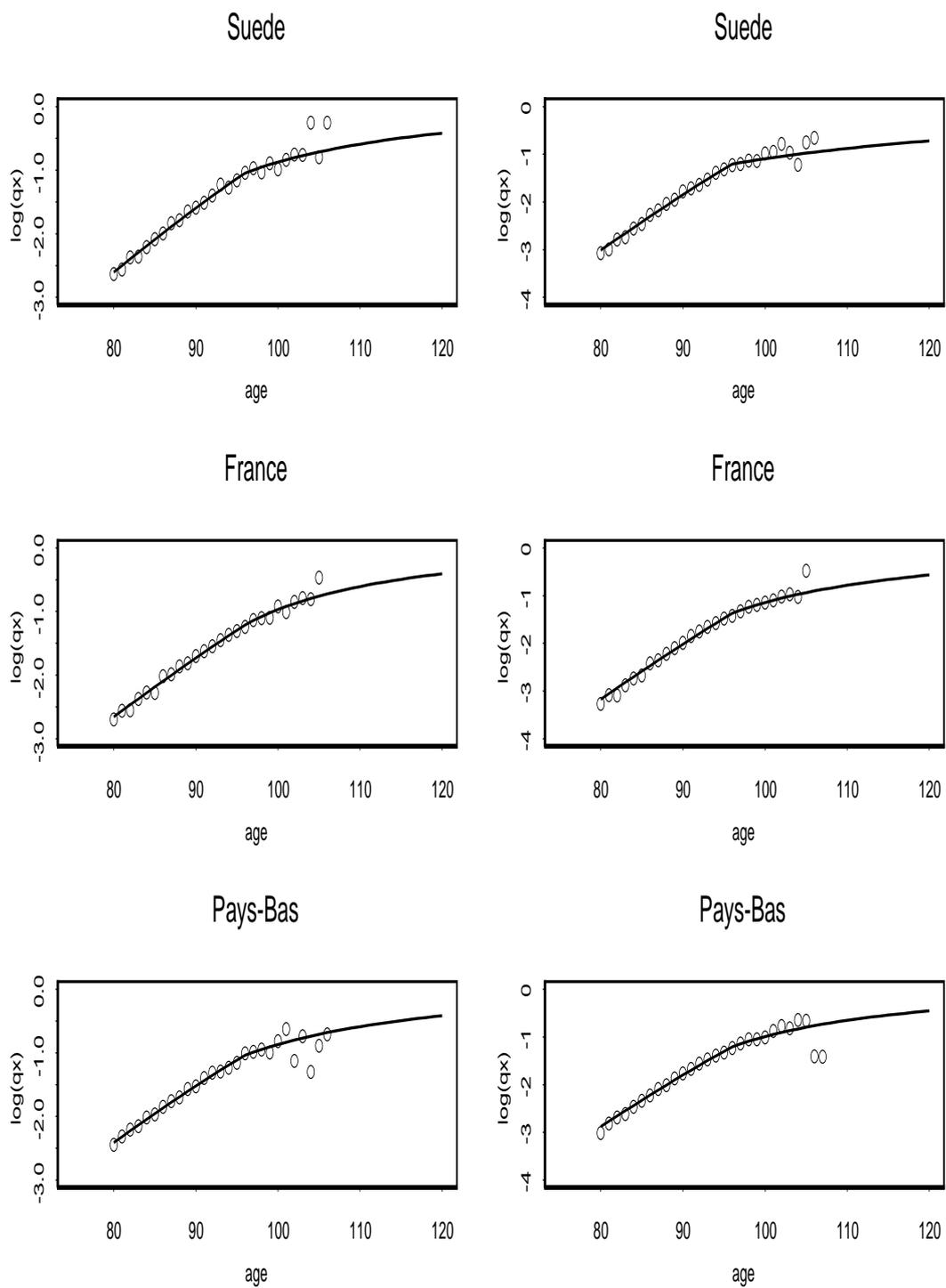


FIG. 2.5 – Comparaison des quotients de mortalité bruts et des quotients de mortalité reconstitués et extrapolés par la méthode de Lindbergson

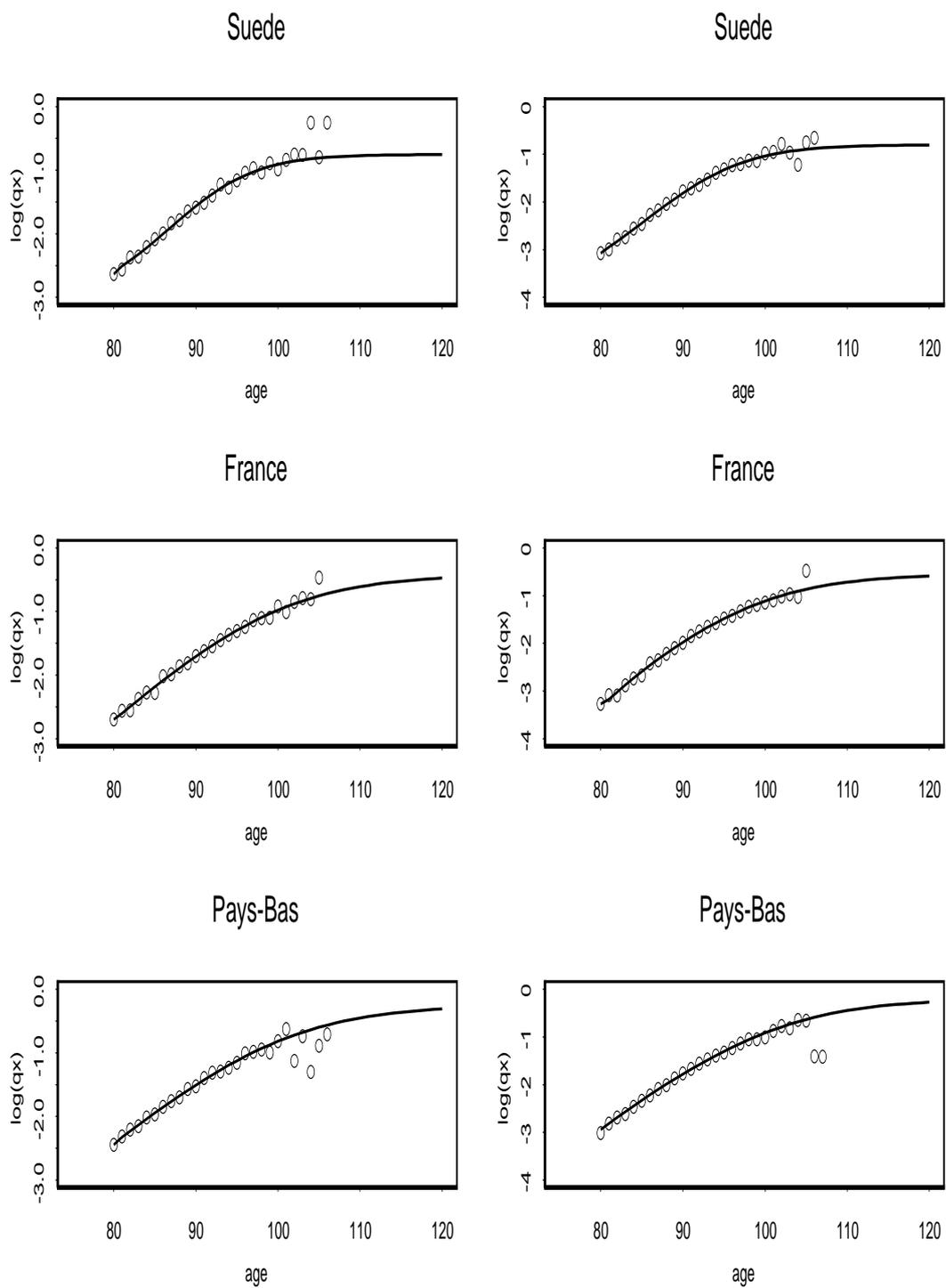


FIG. 2.6 – Comparaison des quotients de mortalité bruts et des quotients de mortalité reconstitués et extrapolés par le modèle Logistique

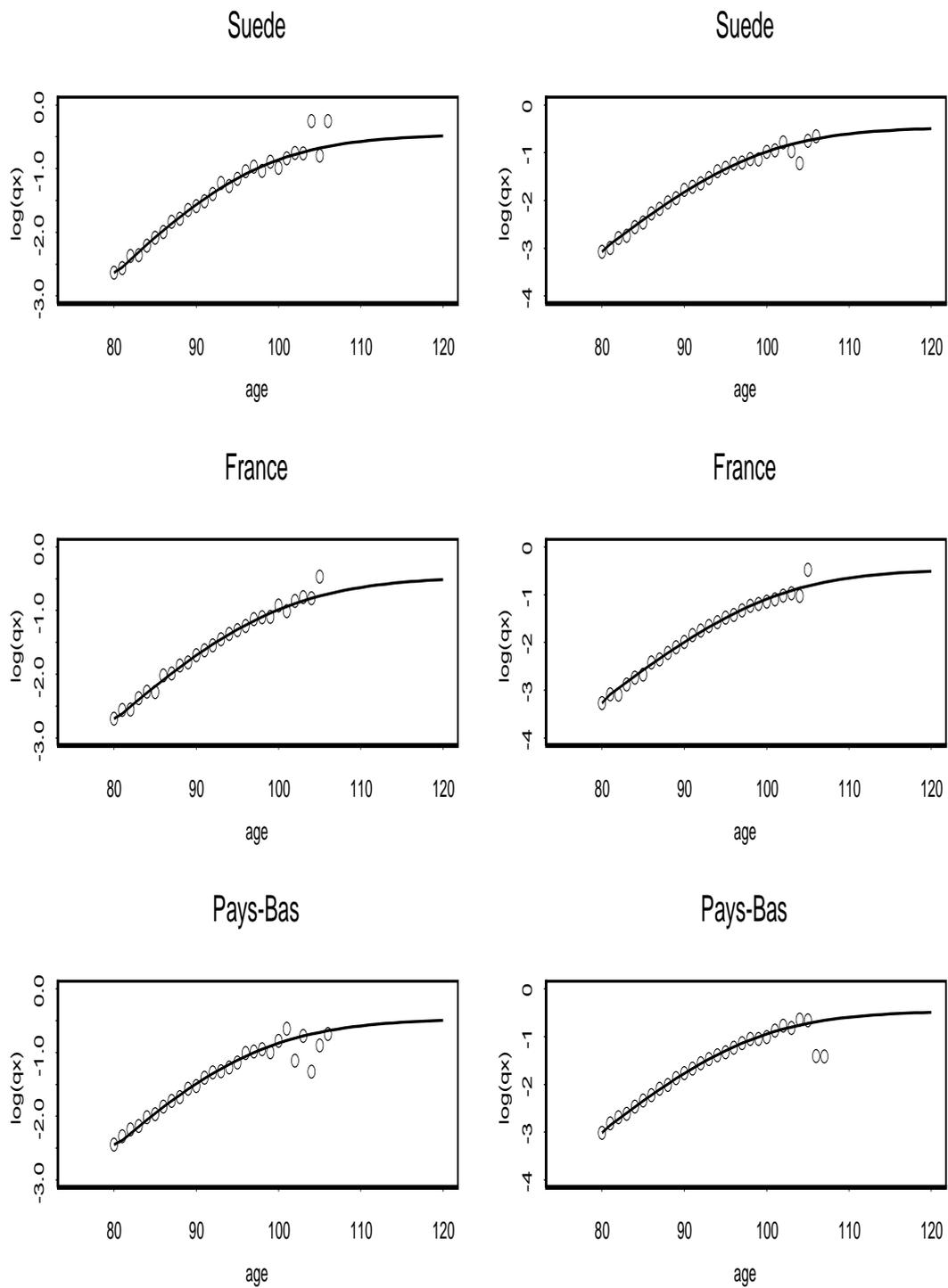


FIG. 2.7 – Comparaison des quotients de mortalité bruts et des quotients de mortalité reconstitués et extrapolés par le modèle de Kannistö

"K", "LG" et "L" désignent respectivement les méthodes de DENUIT & GODERNIAUX, KANNISTÖ, Logistique et de LINDBERGSON).

On voit qu'il y a systématiquement un écart entre d'une part les courbes obtenues par la méthode "DG" et les courbes obtenues par les autres modèles d'extrapolation de la mortalité aux grands âges (Logistique, KANNISTÖ et LINDBERGSON). La courbe obtenue par la méthode "DG" est systématiquement au-dessus des autres courbes. L'âge choisi dans les contraintes (2.2) et (2.3) y est certainement pour quelque chose. Le choix de 120 ans par exemple aurait sans doute donné des courbes similaires à celles obtenues pour les 3 autres méthodes.

3.2 Comparaison des valeurs obtenues pour q_{130}

Dans ce paragraphe, nous allons vérifier si la contrainte $q_{130} = 1$ imposée par DENUIT & GODERNIAUX (2005) est raisonnable du point de vue des 3 autres modèles (LINDBERGSON, KANNISTÖ et Logistique). C'est à la fois un élément de comparaison qu'il est utile d'analyser et en quelque sorte une tentative de rapprochement de 2 classes de méthodes: celles qui proposent une contrainte de fermeture (DENUIT & GODERNIAUX) et celles qui se contentent d'extrapoler les taux de mortalité à partir d'un intervalle d'âges donné (LINDBERGSON, KANNISTÖ et Logistique). Les résultats obtenus sont repris dans le tableau suivant, où on peut lire la valeur de q_{130} selon les différentes méthodes.

Pays	Suède		France		Pays-Bas	
Méthodes	Ho	Fe	Ho	Fe	Ho	Fe
K	0.62733	0.626565	0.622208	0.624114	0.626001	0.627047
LG	0.471319	0.448894	0.648683	0.572185	0.769345	0.904168
L	0.738405	0.756314	0.755114	0.627881	0.739586	0.820532

Force est de constater qu'on est loin de la contrainte $q_{130} = 1$. Cela nous conduit à dire que si les modèles d'extrapolation de la mortalité aux grands âges fournissent des taux très proches des valeurs brutes jusqu'à un certain âge, en revanche elles ne fournissent pas un âge de fermeture de la table raisonnable selon nous dans la mesure où dans notre cas d'étude le cap des 130 ans peut être dépassé avec des probabilités relativement élevées.

3.3 Mesure de l'impact sur la tarification des rentes viagères

La rente viagère est un produit d'assurance vie qui garantit le versement régulier d'une somme d'argent à l'assuré, tant que celui-ci est vivant. La prime unique pure donnant droit à un versement annuel unitaire à terme échu est donnée par

$$a_x = \sum_{k \geq 1} {}_k p_x v^k$$

où $v = (1 + i)^{-1}$ est le facteur d'escompte associé au taux d'intérêt i .

Le Tableau 3.1 donne les résultats obtenus pour le prix de la rente donné par les différentes méthodes en négligeant les âges supérieurs à 120 ans (i.e., en prenant 120 ans comme âge ultime dans les calculs actuariels) et $i = 4\%$. Les chiffres entre parenthèses représentent

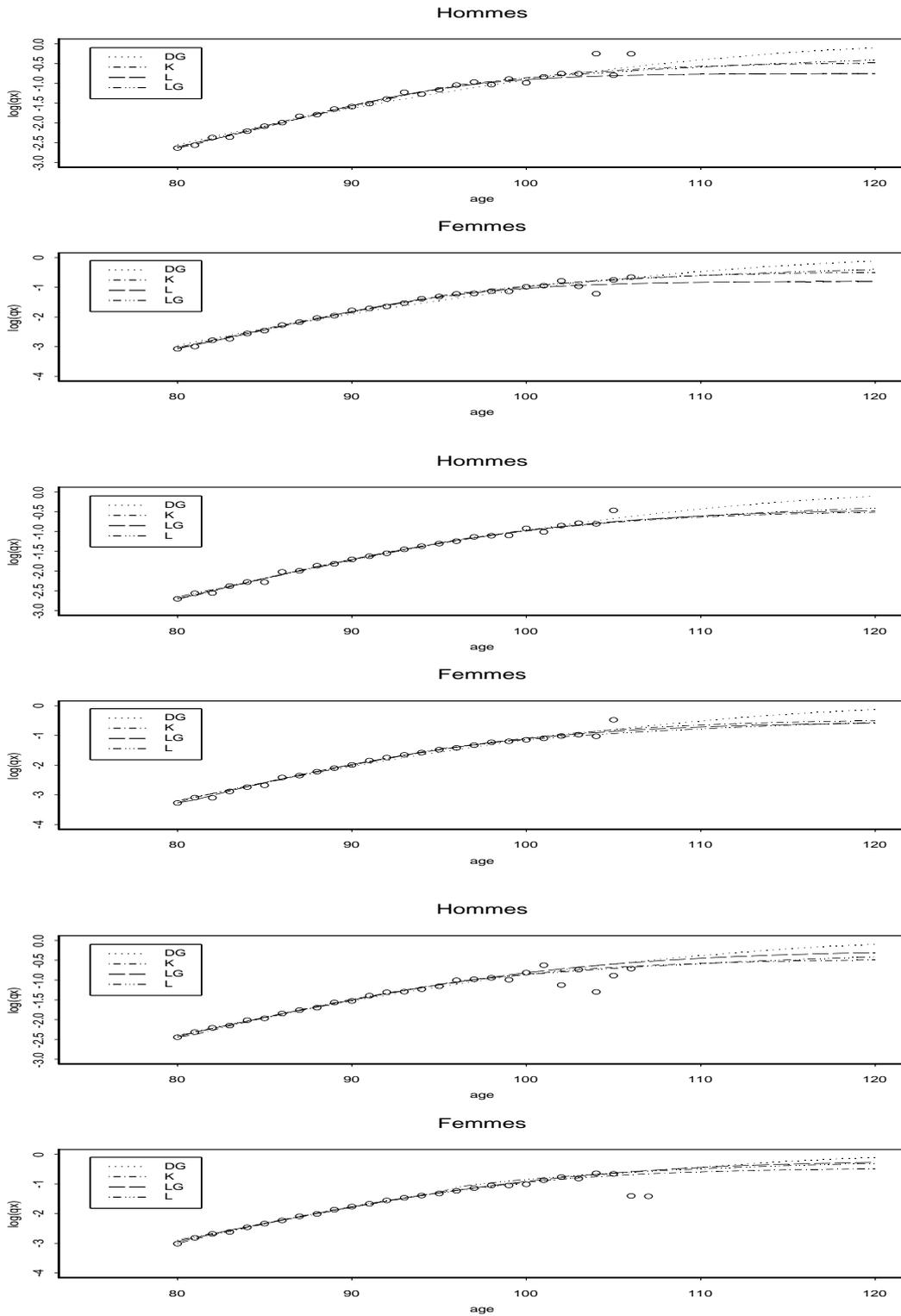


FIG. 3.1 – Comparaison des quotients de mortalité bruts et des quotients de mortalité reconstitués et extrapolés par les méthodes DG, K, LG et L (de haut en bas, Suède, France et Pays-Bas)

l'écart ($a_{65}^{meth} - a_{65}^{Raw}$) entre la rente calculée à partir de la table brute prolongée par l'une ou l'autre méthode et celle calculée sur base de la table brute.

La méthode de DENUIT & GODERNIAUX met clairement en évidence le fait qu'on s'expose à une perte quand on ignore la mortalité aux grands âges dans la tarification des produits tels que la rente viagère. Les modèles de KANNISTÖ et Logistique mettent cet aspect des choses beaucoup moins en évidence. Quant à la méthode de LINDBERGSON, si elle peut apporter des solutions à la connaissance de la mortalité aux grands âges, elle semble inadaptée pour tarifier des produits tels que la rente viagère du fait de l'hypothèse de linéarité des taux aux âges élevés. De toutes les méthodes envisagées, seule celle de DG semble fournir systématiquement un chargement de sécurité par rapport aux tables brutes.

Références

- [1] BOLESZAWSKI, L., & TABEAU, E. (2001). Comparing theoretical age patterns of mortality beyond the age of 80. In *Forecasting Mortality in Developed Countries. Insight from a Statistical, Demographic and Epidemiological Perspective*, pp. 127-155, Kluwer Academic Publishers.
- [2] COALE, A., & GUO, G. (1989). Revised regional model life tables at very low levels of mortality. *Population Index* **55**, 613-643.
- [3] COALE, A., & KISKER, E.E. (1990). Defects in data on old age mortality in the United States: New procedures for calculating approximately accurate mortality schedules and life tables at the highest ages. *Asian and Pacific Population Forum* **4**, 1-31.
- [4] DENUIT, M., & GODERNIAUX, A-C. (2005). Closing and projecting lifetables using log-linear models. Bulletin de l'Association Suisse des Actuaires, sous presse.
- [5] HORIUCHI, S., & WILMOTH, J.R (1998). Deceleration in the age pattern of mortality at older ages. *Demography* **35**, 391-412.
- [6] LINDBERGSON, M. (2001). Mortality among the elderly in Sweden 1988-1997. *Scandinavian Actuarial Journal*, 79-94.
- [7] THATCHER, A.R. (1999). The long-term pattern of adult mortality and the highest attained age. *Journal of the Royal Statistical Society* **162**, 5-43.
- [8] THATCHER, A.R, KANNISTÖ, V. & VAUPEL, J.W. (1998). *The Force of Mortality at Ages 80 to 120*. Odense Monographs on Population Aging **5**: Odense University Press.

Pays	Suède		France		Pays-Bas	
	Ho	Fe	Ho	Fe	Ho	Fe
a_{65}^{Raw}	11.2023	12.7437	11.1441	13.3544	10.2781	12.3251
a_{65}^{DG}	11.2571 (+0.0548)	12.7628 (+0.0191)	11.2271 (+0.083)	13.3824 (+0.028)	10.3234 (+0.0453)	12.3640 (+0.0389)
a_{65}^K	11.2027 (+0.0004)	12.7204 (-0.0233)	11.1493 (+0.0052)	13.3335 (-0.0209)	10.2973 (+0.0192)	12.3375 (+0.0124)
a_{65}^{LG}	11.2082 (+0.0059)	12.7469 (+0.0032)	11.1448 (+0.0007)	13.3564(+0.002)	10.2820 (+0.0039)	12.3233 (-0.0018)
a_{65}^L	11.0688 (-0.1335)	12.6221 (-0.1216)	11.0332 (-0.1109)	13.1252 (-0.2292)	10.1532 (-0.1249)	12.1966 (-0.1285)

TAB. 3.1 – Prime pure de la rente viagère calculée à l'aide des différentes méthodes