

DURÉE DE PERCEPTION DE L'ALLOCATION PERSONNALISÉE D'AUTONOMIE : UN MODÈLE DE PRÉVISION

Clotilde DEBOUT (), Vincent DORTET-BERNADET (**)*

() Direction de la Recherche, des Etudes, de l'Evaluation et des Statistiques, (Drees)
(**) Ensaе*

Introduction

L'Allocation Personnalisée d'Autonomie (APA), entrée en vigueur au 1^{er} janvier 2002, s'adresse aux personnes âgées de 60 ans et plus résidant à domicile ou en établissement et confrontées à des situations de perte d'autonomie. L'APA est une allocation personnalisée répondant aux besoins particuliers de chaque bénéficiaire. Gérée par les départements, elle n'est pas soumise à conditions de ressources, mais son calcul tient compte des revenus des bénéficiaires. Elle permet la prise en charge d'aides et de services diversifiés. Le montant de l'aide dépend de la gravité de la perte d'autonomie. Cette dernière est mesurée à l'aide de la grille nationale AGGIR dont les quatre premiers groupes iso ressources (GIR 1 à 4) ouvrent droit à l'APA.

Depuis 2002 le nombre de bénéficiaires de l'APA augmente de façon importante (+3,8% entre septembre 2007 et septembre 2008). Les projections démographiques, publiées par l'INSEE en 2005 dessinent une augmentation inéluctable de la part des personnes les plus âgées dans les quarante années à venir. L'arrivée aux âges élevés s'accompagne, pour certaines personnes, d'un besoin d'aide pour accomplir les actes essentiels de la vie quotidienne. Même si cette perte d'autonomie n'est pas inéluctable, la forte augmentation à venir du nombre de personnes très âgées fait de la prise en charge des personnes dépendantes et donc de l'évolution du nombre de personnes âgées bénéficiaires de l'APA, un enjeu majeur pour les années futures. Il paraît donc crucial de mieux connaître et de mieux prévoir l'évolution des effectifs de l'APA. Dans ce but, la Drees a mis en place pour un échantillon important de départements (une quarantaine), une remontée des données individuelles sur les bénéficiaires de l'APA. Ces données sont extraites des applications de gestion mises en place par les conseils généraux. Elles contiennent des informations (la date de naissance, l'âge et le parcours dans l'APA) sur toutes les personnes qui ont eu des droits ouverts à l'APA en 2006 ou 2007.

Les données de l'APA ne permettent pas d'avoir beaucoup de recul car le dispositif est relativement récent : les résultats donnés par les modèles de durée classiques sont affectés par le nombre important de données censurées. Ce problème peut être en partie résolu si on considère chaque niveau de dépendance séparément. Comme le temps passé dans un niveau est plus court que le temps total de perception de l'APA il aura moins de chance d'être censuré. Si on analyse aussi les passages entre états on peut estimer une loi qui gouverne le système.

Nous avons choisi d'estimer un modèle de transition comprenant 16 états définis par le croisement des 4 niveaux de dépendance avec les deux types d'hébergement (à domicile ou en établissement) et avec le statut de l'allocataire (soit entrant, soit non entrant, le changement de statut se faisant lors d'un changement de niveau de dépendance). Le modèle retenu généralise le principe de la chaîne de Markov à temps continu en remplaçant les lois exponentielles associées au temps passé dans un état par des lois de Weibull. Ce changement nous permet de prendre en compte des fonctions de risques non constantes pour la durée passée dans chaque état.

A partir des données fournies par trois départements nous proposons une estimation provisoire du modèle qui donne une première idée de la dynamique du système, de la richesse de l'outil mis en œuvre, mais illustre aussi ses limites (hypothèses). Nous montrons que pour un âge fixé, les femmes perçoivent plus longtemps l'allocation que les hommes. En moyenne, sur les trois départements sur

lesquels repose l'étude et d'après les hypothèses retenues, une femme rentrant dans le système en 2007 perçoit l'allocation pendant 4,4 ans contre 2,9 ans en moyenne pour les hommes. L'estimation de modèles de durées différents pour chaque niveau de dépendance et chaque type de lieu de vie permet de fournir des résultats détaillés : on peut par exemple estimer à 3,9 ans le temps moyen passé dans l'APA établissement pour les femmes (contre 2,5 ans pour les hommes). L'estimation du modèle permet aussi de réaliser des projections du nombre de bénéficiaires de l'APA qui s'appuient sur les projections de population par classe d'âge réalisées par l'Insee. Cette partie méritera des approfondissements, suite au problème rencontré pour le passage entre les données observées (passé) et projetées (futur). Ce problème vient à la fois de la base de données qui ne couvre que trois départements (et qui n'est pas suffisamment représentative de la France métropolitaine) et d'un résultat paradoxal de nos projections d'APA pour les années 2008-2009 : elles montrent une baisse peu réaliste de l'effectif des allocataires.

Dans un premier temps nous décrivons plus précisément les données que nous avons traitées et les principales difficultés techniques rencontrées. En particulier nous insistons sur la sélection des données qui influence fortement les estimations de la durée de perception de l'allocation. Nous présentons ensuite les caractéristiques techniques du modèle que nous avons choisi d'estimer avant de présenter, à titre pédagogique et illustratif, les premiers résultats obtenus.

1. Modélisation de la durée de perception de l'APA

Depuis sa création en 2002, le nombre de bénéficiaires de l'APA est en constante progression il est passé de 998 000 fin 2006 à 1 082 000 fin 2007. Cette allocation représente à elle seule plus des quatre cinquième de l'ensemble des aides sociales en faveur des personnes âgées. Dans près de 60% des cas l'allocation aide des personnes vivant à domicile, elle leur permet de rester chez elles même si elles ne peuvent pas accomplir seules certains actes de la vie quotidienne. L'APA est affectée au paiement de dépenses préalablement identifiées dans un plan d'aide par une équipe médico-sociale.

Le montant versé dépend du niveau de dépendance de l'allocataire. Celui-ci est mesuré grâce à la grille AGGIR (cf. encadré ci-dessous). Il s'agit d'une évaluation de l'autonomie des personnes âgées faites à partir de critères comme la capacité à se vêtir, à s'alimenter, à se repérer dans le temps et l'espace. Un algorithme permet de classer les personnes en 6 groupes appelés « groupes iso-ressources » (GIR). Des six groupes seuls les GIR 4 à GIR 1 (le groupe des personnes les plus dépendantes) permettent d'obtenir l'APA.

Tableau 1 : répartition de l'APA par GIR

GIR	APA à Domicile	APA en institution
GIR 4	54,5%	26,0%
GIR 3	22,4%	16,0%
GIR 2	19,8%	43,0%
GIR 1	3,2%	14,8%

Source : Drees, enquête Aide Sociale 2006

Dans 40% des cas l'allocation est versée à des personnes en institution. L'aide sert alors à acquitter une partie du tarif dépendance de l'établissement. Elle est versée par le conseil général soit directement au bénéficiaire, soit aux établissements. La répartition par GIR varie en fonction du type d'hébergement : les résultats du tableau 1 montrent que les GIR4 qui correspondent aux personnes les moins dépendantes sont majoritaires pour l'APA domicile alors que les GIR 2 sont les plus nombreux pour l'APA établissement.

La grille AGGIR (Autonomie gérontologique groupe iso-ressources) classe les personnes âgées en six niveaux de perte d'autonomie :

- GIR 1 : les personnes confinées au lit ou au fauteuil, ayant perdu leur autonomie mentale, corporelle, locomotrice et sociale, qui nécessitent une présence indispensable et continue d'intervenants.

- GIR 2 : les personnes confinées au lit ou au fauteuil dont les fonctions mentales ne sont pas totalement altérées et qui nécessitent une prise en charge pour la plupart des activités de la vie courante, ou celles dont les fonctions mentales sont altérées mais qui ont conservé leurs capacités motrices.

- GIR 3 : les personnes ayant conservé leur autonomie mentale, partiellement leur autonomie locomotrice, mais qui nécessitent quotidiennement et plusieurs fois par jour des aides pour leur autonomie corporelle.

- GIR 4 : les personnes qui n'assument pas seules leur transfert mais qui, une fois levées, peuvent se déplacer à l'intérieur du logement. Elles doivent être aidées pour la toilette et l'habillage.

- GIR 5 : comporte des personnes âgées ayant seulement besoin d'une aide ponctuelle pour la toilette, la préparation des repas et le ménage.

- GIR 6 : réunit les personnes âgées n'ayant pas perdu leur autonomie pour les actes essentiels de la vie courante.

1.1. Caractéristiques des données de l'APA

Pour cette étude nous avons traité les données fournies par trois conseils généraux. Pour chaque département nous avons disposé des informations sur les allocataires qui ont reçu au moins une fois l'APA entre le 1^{er} janvier 2006 et le 31 décembre 2007. Pour chaque individu on connaît sa date de naissance, une évaluation de ses revenus, son sexe, et l'historique de son parcours dans l'APA. Cet historique signale tous les changements de GIR, de type d'hébergement ainsi que la date éventuelle de décès. La base n'est toutefois pas complète car nous ne pouvons pas récupérer des informations sur les personnes en établissement sous-dotation globale. Il est donc impossible de connaître le parcours en établissement pour les personnes passées de domicile à établissement sous dotation globale et l'intégralité de l'historique pour les personnes qui n'ont perçu l'APA qu'en établissement sous dotation globale.

Les données disponibles constituent un stock. L'estimation doit prendre en compte ce mode de sélection car elles ne sont pas représentatives des générations qui chaque année entre dans l'APA. En particulier les personnes qui perçoivent l'APA pendant une durée longue sont surreprésentées. Plus précisément nous avons un problème de troncature à gauche des données : les personnes décédées avant janvier 2006 ne sont pas observées.

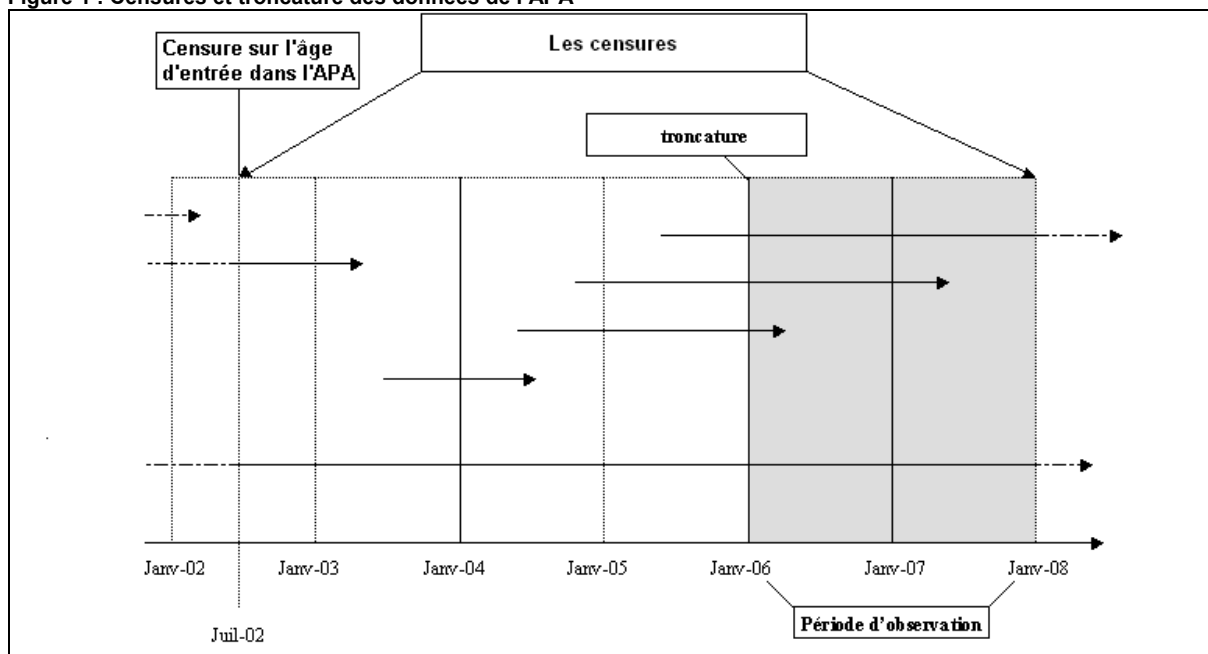
A ce problème de troncature s'ajoutent deux autres problèmes de censure. Nous n'observons pas la totalité de la durée dans l'APA dans les deux cas suivants :

- 1 Pour les personnes qui ont commencé à recevoir l'APA dès le début du dispositif en 2002, on n'observe qu'une partie du temps en dépendance. En effet la plupart d'entre elles auraient perçu l'allocation si elle avait été mise en place plus tôt. On minore donc le temps pendant lesquelles ces personnes auraient pu percevoir l'APA si cette allocation avait existé avant 2002.
- 2 Fin 2007, plus d'un million de personnes âgées sont bénéficiaires de l'APA, personnes pour lesquelles on ne connaît qu'un minorant de leur temps total de perception de l'allocation.

Les modèles que nous allons estimer cherchent principalement à expliquer le temps passé dans l'APA (ou dans un GIR) en fonction de l'âge au moment de l'arrivée dans l'APA (ou l'âge au moment du changement de GIR). La censure en 2002 provoque un deuxième type de problème car elle empêche de connaître l'âge auquel les personnes entrées dans l'APA auraient pu rentrer dans le dispositif s'il

avait existé avant 2002. En fait l'âge que l'on utilise dans ces cas de censure surestime forcément l'âge réel d'entrée en dépendance. Le schéma suivant permet de visualiser les différents problèmes de sélection des données observées en 2006-2007 :

Figure 1 : Censures et troncature des données de l'APA



Avec un modèle paramétrique exponentiel ou de lois de Weibull il est toutefois possible de tenir compte de la troncature à gauche et de la censure à droite des données. On peut obtenir une estimation de la durée totale de perception de l'APA et avec un modèle de hasard proportionnel on peut faire dépendre cette durée de l'âge d'entrée dans le système. Cette approche n'est cependant pas adaptée à nos données : la durée d'observation est très courte et un nombre important de données sont censurées (la majorité des personnes percevant l'APA en 2006-2007 la perçoit encore à la fin de l'année 2007). De plus, avec les historiques des parcours dans le système d'allocation nous disposons de nombreuses données dont on ne tient pas compte en traitant la durée d'allocation en un seul bloc. Ces informations vont pouvoir être utilisées en estimant des modèles de durées pour chaque type de dépendance et en modélisant les transitions entre ces différents états.

Avant de préciser à la section suivante le modèle retenu, nous faisons le point sur les variables disponibles pouvant influencer la durée de perception de l'APA. Avec des modèles semi-paramétriques de Cox nous avons testé l'influence de différents déterminants sur la durée de perception de l'APA : l'âge d'entrée dans le dispositif, le GIR d'entrée, le niveau de ressource, le département et le type d'hébergement que l'on croise avec le niveau de GIR et enfin le zonage en aire urbaine (le ZAU nous permet de distinguer les zones rurales des zones urbaines)¹. Deux paramètres se sont avérés non significatifs : le ZAU et le niveau de ressource. Ce résultat est en partie confirmé par une étude sur les données de l'enquête HID 1998-1999 de l'Insee montrant que la durée de vie en dépendance dépend de l'âge d'entrée mais est peu liée à la catégorie socioprofessionnelle et au fait de vivre en couple (voir [3]). Au final seul l'âge servira à expliquer l'évolution de la dépendance.

Tableau 2 : Estimation de la fonction de risque instantanée de sortie de l'APA par un modèle de Cox

	Hommes	Femmes
Age lors de l'entrée dans l'APA		
De 70 à 75 ans	Réf.	Réf.
De 75 à 80 ans	-0,03 (0,04)	-0,14 (0,04) ***
De 80 à 85 ans	0,11 (0,04) ***	-0,02 (0,03)
De 85 à 90 ans	0,16 (0,04) ***	-0,15 (0,03) ***

¹ Nous avons aussi essayé d'utiliser la variable sur l'état matrimonial : mais dans la plupart des cas il s'agit de l'état matrimonial au moment de l'enquête (et non au début de la dépendance) ce qui biaise le résultat.

De 90 à 95 ans	0,15 (0,05) ***	-0,09 (0,04) **
Plus de 95 ans	0,29 (0,09) ***	0,1 (0,05) *
Département		
Département 1	Réf.	Réf.
Département 2	-0,03 (0,03)	-0,01 (0,02)
Département 3	0,21 (0,05) ***	0,23 (0,04) ***
GIR et type d'hébergement initial		
GIR 4 domicile	Réf.	Réf.
GIR 3 domicile	0,12 (0,04) ***	0,12 (0,03) ***
GIR 2 domicile	0,24 (0,03) **	0,07 (0,03) **
GIR 1 domicile	-0,17 (0,08) ***	-0,32 (0,06) ***
GIR 4 établissement	-0,29 (0,08) ***	-0,18 (0,05) ***
GIR 3 établissement	-0,12 (0,09)	-0,32 (0,07) ***
GIR 2 établissement	-0,1 (0,07)	-0,35 (0,05) ***
GIR 1 établissement	0,44 (0,11) ***	-0,37 (0,07) ***

Source : «Remontées Individuelles APA 2006-2007» Drees pour trois départements

Lecture : Écarts-types entre parenthèses, * : significatif à 10%, ** : significatif à 5%, *** : significatif à 1%

L'effet de l'âge est croissant chez les hommes : le risque instantané de décès croît avec l'âge d'entrée dans l'APA. L'effet est beaucoup moins probant chez les femmes. Il n'y a pas non plus un lien direct entre le niveau de GIR d'entrée et la durée de perception de l'allocation. Nous remarquons toutefois qu'en général à âge donné une personne percevra plus longtemps l'APA si elle est en établissement quand elle la reçoit pour la première fois (à l'exception des GIR 1 établissement pour les hommes). Nous remarquons aussi une différence significative entre les départements : tout étant fixé par ailleurs le risque instantané de décès est plus élevé dans le département 3 (les allocataires perçoivent donc moins longtemps l'APA). Cette différence peut venir notamment de différences socioéconomiques entre les départements, de choix politiques faits en matière de versement de l'allocation, voire de facteurs de santé publique. Ainsi, les décès liés aux diabètes présentent de fortes disparités régionales. De la même façon, les décès liés à une chute accidentelle sont significativement moins fréquents en région parisienne (cf. Disparités régionales de santé, Drees). Cette dépendance importante permet de relativiser la portée des résultats qui vont être présentés : ils n'ont été obtenus qu'à partir des données de trois départements ils ne fournissent donc qu'une idée approximative de la durée de perception de l'allocation pour la France entière.

1.2. Modélisation semi-Markovienne des transitions

Le modèle semi-paramétrique de Cox montre un lien direct entre le niveau GIR d'entrée et la durée totale de perception de l'APA. Plus généralement il semble vraisemblable que le temps passé dans chaque état (obtenu par le croisement des niveaux de GIR et le type d'hébergement) varie en fonction du niveau de dépendance des allocataires. Le temps passé en GIR1 établissement est sans doute plus court que celui passé en GIR4 domicile. Les historiques fournis par les données APA nous permettent de vérifier cette intuition en rendant possible l'estimation d'un modèle de durée pour chaque état de dépendance.

Détailler le temps de perception de l'APA par niveau de dépendance permet aussi d'avoir une idée plus précise du coût de l'allocation. Rappelons en effet, que les sommes versées augmentent quand une personne voit son état de santé s'aggraver. Une prévision de l'évolution des effectifs de chaque GIR et type d'hébergement permettra de mieux évaluer les futures dépenses liées au dispositif.

D'un point de vue statistique, le découpage du temps passé dans le système de l'APA permet d'améliorer la fiabilité des résultats sur au moins deux points. Pour estimer le temps passé dans un état nous sélectionnons toutes les personnes qui ont quitté cet état au cours de la période d'observation : nous avons un échantillon du flux de personnes passant par cet état. Or dans cet échantillon nous retrouvons beaucoup de personnes qui ne sont pas mortes à la fin 2007 mais qui ont tout de même changé de niveau de GIR en 2006-2007 : par rapport à l'approche globale, le nombre d'observations servant à l'estimation du modèle augmente fortement. Cette augmentation est particulièrement importante pour les données de l'APA à domicile : modéliser la durée dans chaque GIR permet en effet de récupérer les informations sur les personnes ayant perçu l'aide à leur domicile mais qui sont parties ensuite dans un établissement placé sous dotation globale.

Le découpage par GIR permet aussi de diminuer l'effet de la censure en 2002 : comme le temps passé dans chaque niveau est plus court que le temps total passé dans l'APA la proportion d'observations censurées diminue. Remarquons que ceci n'est vrai que si l'on considère l'ensemble des états. Pour un niveau particulier comme les GIR1 domicile, nous nous retrouverons avec tout de même peu de données et dans le cas du GIR1 établissement nous observerons une forte proportion de données censurées.

Une façon simple de modéliser les durées passées dans chaque état et les passages d'un état à un autre est de considérer que l'historique de chaque allocataire de l'APA est une réalisation d'une chaîne de Markov à temps continu (les données dont nous disposons nous ont permis d'estimer des modèles à temps continu ce qui différencie notre approche de celle proposées par Olivier Pérou [7] et Michel Duée et al. [4]). Ce modèle suppose que les temps passés dans chaque GIR (et type d'hébergement) sont indépendants. Comme ces durées varient d'un état à l'autre (par exemple le temps passé en GIR4 domicile est en moyenne plus long que celui passé en GIR1 établissement) nous considérons qu'elles suivent des lois différentes que nous faisons varier en fonction de l'âge d'arrivée. D'un point de vue technique, faire dépendre la durée de l'âge d'arrivée dans l'état rend le processus inhomogène car la loi de transition varie en fonction de la date de changement de GIR.

En pratique l'indépendance des durées passées dans chaque état simplifie l'estimation : elle permet une évaluation séparée des paramètres de chaque loi. Pour estimer la loi correspondant au GIR4 domicile nous prenons l'échantillon des durées passées dans cet état par les allocataires : si une personne est passée deux fois dans cet état, les deux durées apparaissent comme deux observations indépendantes dans notre base.

En plus du temps passé dans un GIR il faut aussi déterminer quel sera l'état suivant et avec quelle probabilité. Pour cela nous utilisons des modèles à risque concurrents. Nous souhaitons modéliser la durée T passée dans un état (le GIR4 domicile par exemple) ainsi que l'état de sortie S qui peut prendre M valeurs possibles (notées par commodité $1 \dots M$). Les modèles à risques concurrents suppose qu'il existe M variables aléatoires latentes (non observées) indépendantes T_1, \dots, T_M telles que

$$T = \min(T_1, \dots, T_M) \quad \text{et} \quad S = \arg \min(T_1, \dots, T_M).$$

C'est la cause de changement de GIR la plus rapide à se déclarer qui l'emporte et qui décide de l'état de sortie. Notons $h(t)$ la fonction de hasard de la variable T et $h_m(t)$ la probabilité instantanée de quitter l'état actuel et de sortir vers l'état m , nous avons alors :

$$h(t) = \sum_{m=1}^M h_m(t) \quad \text{et} \quad \lim_{dt \rightarrow 0} P(T \in [t; t + dt], S = j) = h_j(t) \exp\left(-\int_0^t h(u) du\right).$$

La log-vraisemblance du modèle pour les observations $((T_i, S_i))_{i=1}^n$ s'écrit simplement :

$$\begin{aligned} L &= \sum_{i=1}^n \log(h_{S_i}(T_i)) - \int_0^{T_i} h(u) du \\ &= \sum_{m=1}^M \sum_{i=1}^n d_{mi} \cdot \left[\log(h_m(T_i)) - \int_0^{T_i} h_m(u) du \right] - (1 - d_{mi}) \cdot \left[\int_0^{T_i} h_m(u) du \right] \\ &= \sum_{m=1}^M L_m \end{aligned}$$

où d_{mi} est la variable indicatrice valant 1 si l'observation i sort vers l'état m et zéro autrement. La modélisation porte sur les fonctions de hasard $h_m(t)$.

Dans le cas d'une chaîne de Markov, il faut estimer pour chaque couple (état actuel-état de sortie) une loi exponentielle dont le paramètre dépend de l'âge d'arrivée dans le GIR actuel (nous le répartissons en sept classes). Il suffit donc d'utiliser par exemple, la procédure LIFEREG de Sas pour chaque sortie possible en conditionnant par les variables dichotomiques des classes d'âge. Cette approche s'avère cependant trop simple pour modéliser les parcours dans l'APA. L'utilisation d'un modèle de chaîne de Markov implique que le processus est sans « mémoire », la loi du temps restant à passer dans un état est indépendante du temps que l'on y a déjà passé. Nous montrons que cette hypothèse n'est pas raisonnable en comparant les résultats de deux régressions :

- la régression en fonction de l'âge du temps total passé par les femmes en GIR4 domicile
- la régression en fonction de l'âge du temps passé par les femmes en GIR4 quand elles ont passé déjà un an dans cet état

Le tableau suivant montre que les deux régressions ont des résultats significativement différents :

Tableau 3 : Comparaison de deux régressions sur le temps passé en GIR 4 domicile

Classe d'âge	Régression de la durée totale en GIR4 domicile		Régression de la durée restante en GIR4 domicile après un an dans cet état	
	Estimateur obtenu par la procédure LIFEREG	Ecart-type de l'estimateur	Estimateur obtenu par la procédure LIFEREG	Ecart-type de l'estimateur
Moins de 70 ans	2,41	0,08	1,98	0,08
70 à 74 ans	2,44	0,07	2,01	0,07
75 à 79 ans	2,47	0,05	2,03	0,05
80 à 84 ans	2,72	0,05	2,24	0,05
85 à 89 ans	2,67	0,06	2,14	0,06
90 à 94 ans	2,75	0,09	2,22	0,09
95 ans et plus	3,18	0,27	2,52	0,27

Source : « Remontées Individuelles APA 2006-2007 » Drees pour trois départements

Pour tenir compte de ce phénomène nous avons décidé de généraliser le modèle en utilisant des lois de Weibull. Les lois de Weibull constituent une généralisation naturelle des lois exponentielles associées aux chaînes de Markov. Elles permettent de décrire des phénomènes où la fonction de hasard n'est pas constante : la probabilité instantanée de passage vers l'une des sorties possibles devenant plus ou moins importante avec le temps. Dans les estimations que nous avons menées nous avons testé pour chaque état et chaque sortie l'hypothèse nulle de la loi exponentielle : les résultats sont variés, certains états acceptant l'hypothèse markovienne même si la majorité correspondent à des lois de Weibull dont le paramètre de dispersion est différent de 1.

Dans le cas où l'on suppose qu'il n'y a pas de paramètres communs aux lois de sortie vers les différents états m alors la maximisation de la vraisemblance est simple car les fonctions L_m peuvent être maximisées séparément. Si on examine l'écriture d'une fonction L_m on voit que sa maximisation revient à estimer un modèle de durée à partir des observations T_i qui sont supposées censurées lorsque la sortie n'est pas l'état m . L'utilisation de lois de Weibull oblige cependant à un traitement spécial de la censure en 2002.

La mise en place du dispositif en 2002 représente une censure à droite car la durée n'est pas entièrement observée. Le problème est que cette censure ne peut être interprétée comme une sortie particulière comme dans le modèle à risques concurrents de base. Nous avons donc été obligés de calculer à la main l'estimateur du maximum de vraisemblance avec un modèle de Weibull où le paramètre de dispersion est commun à toutes les sorties de l'état. Le cas plus complexe où l'on voudrait considérer des paramètres de dispersion différents pour chaque sortie de l'état n'est pas aussi facilement calculable. Nous nous sommes donc imposé cette restriction pour les états d'entrée dans le dispositif (on verra à la section suivante que l'on distingue les états d'entrées des états en cours de traitement). Pour présenter ce calcul nous complétons nos notations en définissant :

- k_{mi} : l'indicatrice valant 1 quand l'individu i sort vers m et que sa durée n'est pas censurée,

- l_{mi} : l'indicatrice valant 1 quand l'individu i sort vers m et que sa durée est censurée,
- $\bar{F}(t) = P(T > t) = \exp\left(-\int_0^t h(u)du\right)$, la fonction de survie associée à T .

La log-vraisemblance pour les observations $T_i (i=1 \dots n)$ s'écrit :

$$L = \sum_{m=1}^M \sum_{i=1}^n k_{mi} \cdot \ln(h_m(T_i)) - \int_0^{T_i} h_m(u)du + l_{mi} \cdot \ln\left[\int_{T_i}^{+\infty} h_m(u) \bar{F}(u) du\right].$$

Pour le modèle de Weibull avec un unique paramètre de dispersion α , la fonction de hasard vers la sortie m s'écrit (en notant $a(i)$ la classe d'âge de l'individu i) :

$$h_m(T_i) = \alpha T_i^\alpha \lambda_{a(i)}^m.$$

On dira alors que i appartient à l'ensemble C quand l'observation est censurée. La log-vraisemblance devient alors :

$$L = \sum_i \left(\ln[\lambda_{a(i)}] - T_i^\alpha \sum_m \lambda_{a(i)}^m \right) + \sum_{i \in C^c} \ln(\alpha T_i^\alpha) - \sum_a N_{censuré}^a \ln\left[\sum_m \lambda_a^m\right]$$

où $N_{censuré}^a$ est le nombre de censures dans la classe a . L'expression de $\hat{\lambda}$ en fonction de $\hat{\alpha}$ est :

$$\hat{\lambda}_a = \frac{N_m^a}{\sum_{i:a(i)=a} T_i^{\hat{\alpha}}} \cdot \frac{N^a - N_{censuré}^a}{N^a}$$

où N_m^a est le nombre d'individus de la classe a sortant vers m et N^a le nombre total d'observations dans la classe d'âge a . Pour le paramètre de dispersion le maximum de vraisemblance s'obtient en annulant la fonction :

$$G(\alpha) = \frac{N_{non\ censuré}}{\alpha} + \sum_i \ln(T_i) T_i^\alpha \cdot \frac{N^{a(i)} - N_{censuré}^{a(i)}}{\sum_{i:a(i)=a} T_i^\alpha} + \sum_{i \in C^c} \ln(T_i)$$

Une valeur approchée de $\hat{\alpha}$ est obtenue à l'aide de la méthode de Newton.

1.3. Choix du modèle

Rappelons tout d'abord les objectifs que nous devons remplir pour mener à bien la modélisation :

- Nous voulons pouvoir évaluer le temps passé par un allocataire dans chaque GIR et type d'hébergement.
- Pour chaque individu nous souhaitons connaître à chaque instant les probabilités de changement de niveau de dépendance.
- Enfin il faut disposer d'un modèle propre à chaque sexe et l'âge doit servir de variable explicative des durées observées.

Pour atteindre ces objectifs nous avons procédé par étapes, le but étant prendre en compte le maximum d'information fournies par les données tout en limitant la complexité du modèle.

1.3.1. Choix des transitions significatives

Sachant qu'il y a quatre niveaux de GIR et deux types d'hébergement (soit 8 états) cela fait a priori soixante quatre transitions à modéliser en comptant bien sur le décès comme le neuvième état possible. En fait, le nombre de transitions à modéliser est plus faible, seules certaines transitions étant suffisamment courantes pour être retenues par le modèle (cf. tableau 4) :

Tableau 4 : Matrice de transition entre les états de l'APA

	GIR4 domicile	GIR3 domicile	GIR2 domicile	GIR1 domicile	GIR4 établ.	GIR3 établ.	GIR2 établ.	GIR1 établ.	Sortie du système
GIR4 domicile		34%	20%	1%	12%	1%	1%	0%	31%
GIR3 domicile	7%		35%	3%	1%	15%	2%	1%	37%
GIR2 domicile	3%	4%		9%	1%	1%	21%	1%	61%
GIR1 domicile	1%	1%	5%		0%	0%	1%	16%	76%
GIR4 établissement	0%	0%	0%	0%		27%	35%	6%	32%
GIR3 établissement	0%	0%	0%	0%	15%		46%	10%	28%
GIR2 établissement	0%	0%	0%	0%	6%	11%		30%	52%
GIR1 établissement	0%	0%	0%	0%	0%	1%	14%		85%

Source : «Remontées Individuelles APA 2006-2007» Drees pour trois départements
Lecture : pourcentage des départs d'un état (en ligne) vers une sortie (en colonne)

Ces résultats nous dispensent de tenir compte des retours à domicile car ils sont dans les faits trop rares pour être modélisés. Nous avons signalé en introduction que les passages de domicile à établissement n'étaient pas connus dès que l'établissement d'accueil est sous dotation globale (c'est le cas de la majorité des institutions du département 2). Nous savons seulement que l'allocataire part en institution sans connaître son nouveau GIR. Les résultats obtenus à partir des établissements qui ne sont pas placés sous dotations globale montrent que dans la plupart des cas ces changements se font à niveau de GIR constant. En supposant que le parcours des personnes en établissements hors dotation globale soit le même que celui des personnes en établissements sous dotation globale, nous pouvons supposer que lors d'une transition domicile-établissement le GIR ne change pas. Ceci permet de conserver toutes les observations collectées pour estimer la loi du système dans la partie domicile. Le départ du domicile le parcours reste toujours inconnu pour les observations avec dotation globale et elles ne peuvent rien nous apprendre sur la partie établissement. Au final nous ne modéliserons que les transitions qui représentent au moins 4% des changements (soit 29 transitions au lieu de 64). Ainsi pour sortir de l'état GIR4 domicile on considèrera que seules sont possibles les transitions vers GIR3 et GIR2 domicile, vers GIR4 établissement et enfin la sortie du système d'allocation qui correspond le plus souvent à un décès.

1.3.2. Etat d'entrée versus état en cours d'APA

Une régression montre qu'à âge fixé, la durée de vie dans un état augmente si cet état est celui d'arrivée dans l'APA. Cette propriété n'est pas vérifiée si on considère que la loi du système est celle d'une chaîne de Markov et même si on remplace les lois exponentielles par des lois de Weibull. Ce phénomène peut-il s'interpréter comme la conséquence de l'aggravation de l'état de santé? Une personne commençant à percevoir l'APA en GIR4 établissement n'est peut-être pas semblable à un autre allocataire du GIR4 établissement qui percevait déjà l'APA quand elle vivait à domicile. Nous pouvons aussi voir ici une conséquence de la censure en 2002. En effet seules les observations correspondant à des entrées dans l'APA sont censurées, or l'âge d'entrée dans un état est surestimé pour une observation censurée. Comme la durée de vie décroît avec l'âge on aurait bien pour un même âge d'entrée mesuré, des espérances supérieures pour les entrants dans l'APA à celles des personnes qui ne font que changer de niveau de dépendance.

Pour s'adapter à cette propriété des données nous avons décidé de considérer deux 'sous-états' différents pour chaque couple (niveau de GIR- type d'hébergement). A priori on ne perd rien par rapport à la modélisation initiale : si la distinction que nous venons d'introduire est fictive les estimations donneront juste des résultats très semblables sur la loi associée aux deux sous-états.

Nous avons vu que le problème de la censure empêchait d'estimer facilement des modèles avec des paramètres de dispersion propres à chaque sortie. Or le découpage supprime la censure pour les sous-états qui correspondent à des changements de GIR. Il a donc rendu possible l'estimation du modèle général pour ces transitions. Nous pouvons procéder à une estimation à la main qui comprend donc un unique paramètre de dispersion pour les sous-états d'entrée dans l'APA (de façon à tenir compte de la censure en 2002) et autant de paramètres de dispersion que de sorties pour les sous-états de changement de GIR.

1.4. Choix des données

Nous avons calculé différentes versions du modèle en faisant varier plusieurs paramètres :

- Le choix des données utilisées : nous avons utilisé la base des changements d'état au cours de l'année 2006, celle de l'année 2007 et enfin la fusion de ces deux bases.
- La méthode d'estimation : nous avons estimé un modèle complet à la main en regroupant éventuellement les classes d'âge de façon à ce que les paramètres soient tous significatifs pour chaque couple (état-sortie). Nous avons aussi mis en place une macro Sas permettant d'estimer de façon automatique toutes les transitions à partir d'un choix fixe de régresseurs. Pour la méthode automatique nous avons choisi de n'utiliser que quatre classes d'âge alors que nous avons pu tester jusqu'à sept classes d'âge pour l'estimation à la main.
- Enfin pour le modèle estimé à la main nous avons réservé un traitement particulier pour les personnes rentrées directement en GIR1 établissement. En effet pour ces profils relativement rares on observe un nombre étonnant de données censurées. Nous avons choisi de ne plus tenir compte de la censure afin de montrer la forte influence quelle a sur les résultats obtenus.

Au total nous retenons 4 versions différentes : un modèle estimé à la main sur les données 2006-2007 et trois modèles estimés par la macro sur les données 2006, 2007 et 2006-2007. Pour comparer les différentes versions nous présentons dans le tableau suivant les résultats obtenus pour les femmes entrant en GIR4 domicile et celles qui viennent de changer de GIR et arrivent en GIR4 établissement.

Tableau 5 : Comparaison des durées moyennes (en année) estimées à partir de différentes spécifications du modèle

	Age d'entrée dans l'état	Modèle sélectionné à la main sur les données 2006-2007	Modèle estimé par la macro Sas sur les données 2006-2007	Modèle estimé par la macro Sas sur les données 2006	Modèle estimé par la macro Sas sur les données 2007
Durée totale de perception de l'APA lors d'une entrée en GIR 4 domicile	60 ans	5,2	5,4	4,8	5,5
	65 ans	5,1	5,3	4,8	5,4
	70 ans	5,1	5,2	4,8	5,4
	75 ans	5,1	5,1	4,7	5,2
	80 ans	4,6	4,6	4,3	4,6
	85 ans	4,3	4,3	4,2	4,5
	90 ans	4,2	4	3,9	4,1
	95 ans	3,4	4,1	3,9	4,1
Durée restante de perception de l'APA lors d'un changement pour GIR 4 établissement	60 ans	4,3	5	4,4	4,3
	65 ans	4,3	4,8	4,3	4,3
	70 ans	4,2	3,8	4,2	4,2
	75 ans	4	3	4,1	4,1
	80 ans	3	2,6	2,9	2,8
	85 ans	2,6	2,8	2,9	2,7

	90 ans	2,9	2,7	2,6	2,9
	95 ans	2,1	4,8	2,7	2,9

Source : «Remontées Individuelles APA 2006-2007» Drees pour trois départements

Le choix des données à une influence sur les résultats : si on utilise seulement les données les plus anciennes la durée de perception semble être sous-estimée. C'est particulièrement vrai pour l'APA à domicile cela semble moins vrai en institution. Ce résultat est plutôt inquiétant car il semble prouver que le système n'est pas stationnaire (les politiques d'attribution de l'allocation auraient changé au cours du temps?) ou alors que l'état de santé de la population à niveau de dépendance constant va en s'améliorant. Ces facteurs n'ont pas été pris en compte dans nos estimations ce qui nuit sans doute à leur pouvoir prédictif. La méthode d'estimation semble aussi avoir une influence : on la perçoit surtout pour les résultats sur les femmes qui changent de GIR car pour ces états on a pu sélectionner des paramètres de dispersion propre à chaque sortie. L'effet est cependant moins important que celui dû aux choix des données.

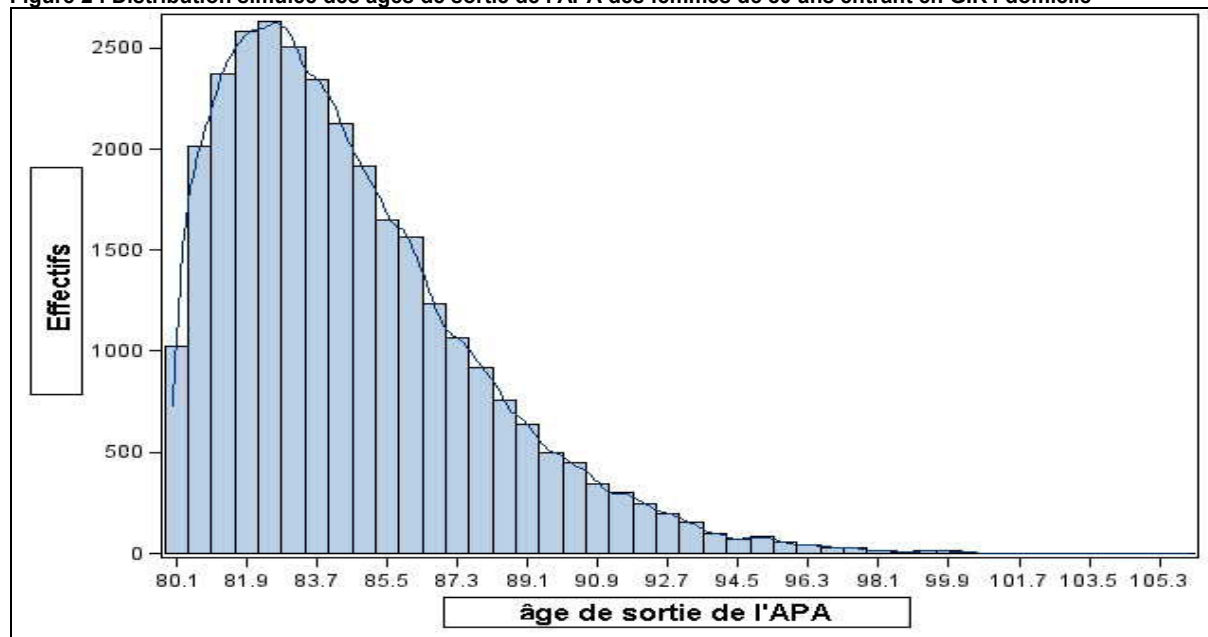
2. Résultats

Nous décidons finalement de présenter en détail les résultats obtenus par la méthode automatique sur les données les plus récentes car ce sont celles qui a priori nous fournissent le plus d'information sur l'évolution future du système et celles aussi qui sont le moins influencées par la censure en 2002. Il aurait été plus logique de présenter un modèle estimé à la main sur les données récentes de 2007 mais faute de temps nous n'avons pas pu le faire. Remarquons tout de même dans le tableau 5 les faibles écarts entre les résultats à la main sur 2006-2007 et les résultats de la macro Sas sur 2007 : nous n'obtiendrons sans doute pas un résultat très différent en estimant le modèle à la main sur les données 2007.

2.1. La durée de perception de l'APA

L'estimation du modèle permet tout d'abord de calculer la durée moyenne de perception de l'allocation pour une personne rentrant dans l'APA ou plus généralement pour une personne qui change de niveau de dépendance. Les tableaux 7 et 8 mettent en évidence l'effet de l'âge et du niveau de dépendance sur la durée de perception de l'APA pour les femmes. Ils permettent aussi de montrer la nette différence entre la durée de perception de l'APA d'une personne qui rentre dans l'APA et celle beaucoup plus faible d'une personne qui atteint le même niveau de dépendance mais qui perçoit déjà l'allocation. Ceci justifie le découpage de chaque état en deux sous-états. Ces chiffres ont été obtenus par simulation de parcours fictifs dans l'APA gouvernés par les lois que nous venons d'estimer. Remarquons que la loi de la durée de vie en dépendance est compliquée. Pour chaque réalisation nous devons faire la somme de plusieurs variables aléatoires qui correspondent aux temps passés dans chacun des états possibles. Chaque variable aléatoire suit une loi de Weibull dont les paramètres dépendent de l'état et de l'âge au moment de l'entrée dans l'état. Les simulations nous permettent d'obtenir des distributions de durée de vie dans l'APA. Par exemple pour les femmes de 80 ans entrant dans l'APA en GIR4 domicile on obtient le graphe suivant :

Figure 2 : Distribution simulée des âges de sortie de l'APA des femmes de 80 ans entrant en GIR4 domicile



Source : Simulations de 30000 parcours à partir du modèle estimé sur les transitions de 2007

Notons que la durée moyenne obtenue pour les personnes entrant directement en Gir1 ou 2 en établissement est plus élevée que ce qui est attendu. On obtient un résultat qui ne suit pas la logique constatée jusqu'alors : la durée de perception de l'allocation décroît en fonction de la gravité de la dépendance. Ces résultats sont dus à une forte proportion de données censurées. Comme le montre le tableau suivant, près de 40% des personnes qui sont rentrées dans l'APA en GIR 1 et qui sont sorties de cet état en 2006 ou 2007 sont rentrées au tout début du dispositif. La censure est nettement plus importante que pour les autres niveaux de dépendance :

Tableau 6 : Part des données censurées dans les changements de GIR en 2006 et 2007

GIR d'entrée	GIR 4 domicile	GIR 3 domicile	GIR 2 domicile	GIR 1 domicile	GIR 4 étab.	GIR 3 étab.	GIR 2 étab.	GIR 1 étab.
Non censuré	83%	85%	83%	75%	84%	88%	75%	61%
Censuré	17%	15%	17%	25%	16%	12%	25%	39%

Source : «Remontées Individuelles APA 2006-2007» Drees pour trois départements

Le niveau de censure varie en fonction de l'état de la sortie de GIR 1 établissement : il est de 18% pour les personnes passant en GIR 2 établissement (ce qui est comparable aux autres GIR) et de 42% pour les personnes décédées en 2006 ou 2007. Les personnes entrant en GIR 1 établissement constituent donc une population hétérogène comprenant sans doute des personnes fortement dépendantes (handicapées avant de rentrer dans l'APA ?) mais dont l'état de santé est stable et qui perçoivent l'APA longtemps, et une autre population dont l'état de santé s'aggrave à un rythme proche de celui mesuré pour les autres allocataires de l'APA. En l'absence d'information supplémentaire sur la nature de la dépendance on ne peut pas faire la différence entre ces deux sous-populations. Comme seulement 1% des personnes percevant l'APA sont des entrants directs en GIR1 établissement nous avons laissé ce problème de côté.

Résultats du modèle de transition :

Espérance de vie des femmes percevant l'APA à domicile, estimation provisoire à partir de 3 départements

Tableau 7 : Durée moyenne (en années) de perception de l'APA pour les femmes percevant l'APA à domicile en 2007

Age	GIR actuel	Durée totale de perception de l'APA lors de l'entrée dans le système	Écart-type	Durée restante de perception de l'APA lors d'un changement de GIR	Écart-type
60 ans	GIR 4	5,5	4,1	4,0	3,3
65 ans	GIR 4	5,4	4,0	4,0	3,4
70 ans	GIR 4	5,4	4,0	4,0	3,3
75 ans	GIR 4	5,2	3,7	3,9	3,1
80 ans	GIR 4	4,6	3,3	3,3	2,7
85 ans	GIR 4	4,5	3,2	3,3	2,7
90 ans	GIR 4	4,1	3,0	3,1	2,5
95 ans	GIR 4	4,1	3,0	3,2	2,5
60 ans	GIR 3	4,9	4,0	3,6	3,1
65 ans	GIR 3	4,9	3,9	3,6	3,1
70 ans	GIR 3	4,8	3,9	3,5	3,0
75 ans	GIR 3	4,7	3,7	3,5	2,9
80 ans	GIR 3	4,1	3,3	2,9	2,5
85 ans	GIR 3	4,0	3,2	2,9	2,4
90 ans	GIR 3	3,5	2,9	2,9	2,3
95 ans	GIR 3	3,5	2,9	2,9	2,3
60 ans	GIR 2	4,3	4,3	2,6	2,7
65 ans	GIR 2	4,3	4,2	2,6	2,7
70 ans	GIR 2	4,3	4,1	2,6	2,6
75 ans	GIR 2	4,3	4,1	2,6	2,5
80 ans	GIR 2	3,7	3,6	2,4	2,2
85 ans	GIR 2	3,7	3,6	2,4	2,2
90 ans	GIR 2	3,3	3,2	1,9	1,9
95 ans	GIR 2	3,2	3,3	2,0	1,9
60 ans	GIR 1	4,7	6,3	2,2	2,3
65 ans	GIR 1	4,7	6,2	2,2	2,4
70 ans	GIR 1	4,7	6,2	2,2	2,3
75 ans	GIR 1	4,6	6,1	2,1	2,2
80 ans	GIR 1	3,9	5,3	1,7	1,8
85 ans	GIR 1	3,8	5,2	1,7	1,8
90 ans	GIR 1	4,7	6,6	1,3	1,4
95 ans	GIR 1	4,8	6,9	1,3	1,4

Source : «Remontées Individuelles APA 2006-2007» Drees pour trois départements

Lecture : Les colonnes 'durée' donnent la moyenne des lois estimées des durées de vie, les colonnes écart-type donnent l'écart-type de ces lois.

Résultats du modèle de transition :

Espérance de vie des femmes percevant l'APA en établissement, estimation provisoire à partir de 3 départements

Tableau 8 : Durée moyenne (en années) de l'allocation pour les femmes percevant l'APA en établissement

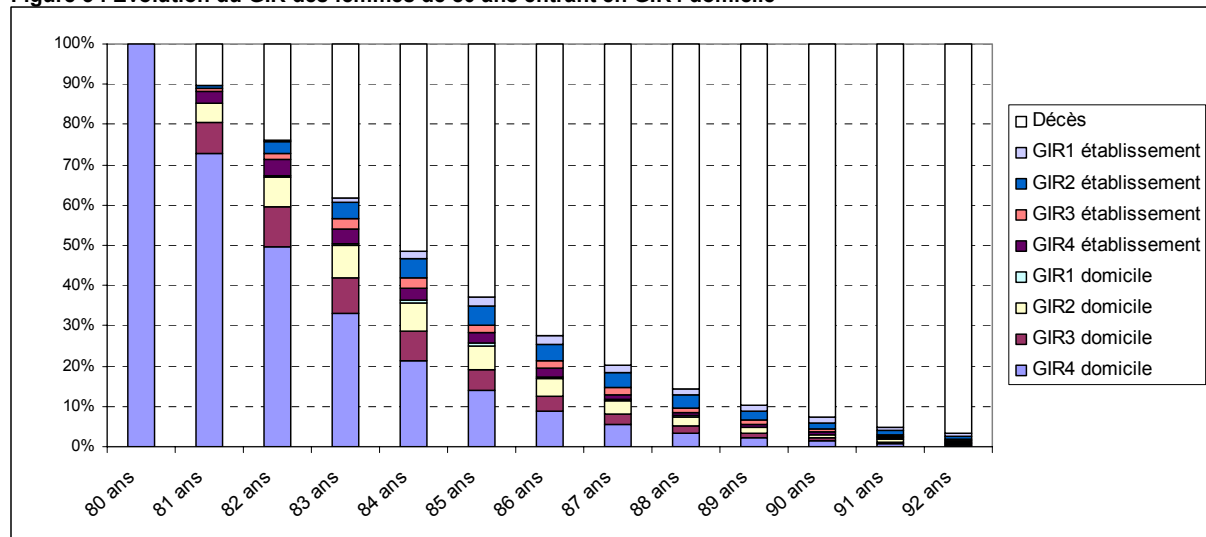
Age	GIR actuel	Temps moyen de perception de l'APA lors de l'entrée dans le système	Écart-type	Durée restante de perception de l'APA lors d'un changement de GIR	Écart-type
60 ans	GIR 4	5,9	4,1	4,3	3,6
65 ans	GIR 4	5,9	4,1	4,3	3,5
70 ans	GIR 4	5,8	3,9	4,2	3,4
75 ans	GIR 4	5,6	3,6	4,1	3,2
80 ans	GIR 4	3,7	2,7	2,8	2,5
85 ans	GIR 4	3,7	2,7	2,7	2,4
90 ans	GIR 4	3,5	2,6	2,9	2,5
95 ans	GIR 4	3,5	2,5	2,9	2,5
60 ans	GIR 3	5,1	4,0	4,1	3,3
65 ans	GIR 3	5,2	3,9	4,1	3,3
70 ans	GIR 3	5,1	3,8	4,1	3,2
75 ans	GIR 3	4,9	3,5	3,9	3,0
80 ans	GIR 3	3,2	2,6	2,9	2,4
85 ans	GIR 3	3,2	2,5	2,9	2,4
90 ans	GIR 3	3,3	2,6	2,6	2,2
95 ans	GIR 3	3,3	2,6	2,6	2,2
60 ans	GIR 2	6,1	5,3	3,1	3,0
65 ans	GIR 2	6,1	5,3	3,2	3,0
70 ans	GIR 2	6,0	5,1	3,1	3,0
75 ans	GIR 2	6,0	5,1	3,1	2,7
80 ans	GIR 2	4,0	3,5	2,4	2,2
85 ans	GIR 2	4,0	3,4	2,4	2,2
90 ans	GIR 2	3,1	2,8	2,2	2,0
95 ans	GIR 2	3,1	2,8	2,1	2,0
60 ans	GIR 1	9,6	15,0	2,2	2,4
65 ans	GIR 1	9,4	15,1	2,2	2,4
70 ans	GIR 1	9,5	15,1	2,2	2,3
75 ans	GIR 1	9,5	15,3	2,2	2,2
80 ans	GIR 1	7,1	11,5	1,9	1,9
85 ans	GIR 1	7,0	11,4	1,9	1,8
90 ans	GIR 1	6,1	10,0	1,7	1,6
95 ans	GIR 1	6,1	9,9	1,6	1,6

Source : «Remontées Individuelles APA 2006-2007» Drees pour trois départements

Lecture : Les colonnes 'durée' donnent la moyenne des lois estimées des durées de vie, les colonnes écart-type donnent l'écart-type de ces lois.

Nous pouvons aussi simuler à partir du modèle l'évolution par GIR et par type d'hébergement : elle montre la décroissance rapide du nombre d'allocataires et la part plus importante des personnes en établissement au fur et à mesure des années.

Figure 3 : Évolution du GIR des femmes de 80 ans entrant en GIR4 domicile



Source : Simulations de 30000 parcours à partir du modèle estimé sur les transitions de 2007

2.2. Projections démographiques : méthodologie et difficultés rencontrées

Dans cette partie nous présentons la démarche que nous avons retenue pour réaliser des projections de bénéficiaires de l'APA. Le travail s'est déroulé en trois temps. Nous avons tout d'abord calculé la répartition par classe d'âge et GIR d'entrée des personnes qui sont entrées dans l'APA en 2007. Nous avons ainsi pu simuler leur espérance de vie et leur parcours futur dans l'APA. La deuxième étape se concentre sur le calcul de l'évolution des personnes qui percevaient l'APA à la fin de 2007 : la difficulté vient du fait que cette date ne correspond pas forcément à un changement de GIR. Il faut alors estimer la loi conditionnelle du temps restant dans l'état actuel sachant le temps que l'on y a déjà passé. A l'aide de simulations et d'une approximation (une régression par un polynôme d'ordre deux) nous obtenons une expression calculable de cette loi conditionnelle. Enfin pour réaliser les projections de bénéficiaires de l'APA nous utilisons les projections démographiques de l'Insee sur les évolutions par classe d'âge sur les quarante prochaines années. En conservant les taux d'incidence (c'est à dire la probabilité à âge fixé qu'une personne rentre dans l'APA) estimés pour l'année 2007, nous pouvons ajouter chaque année une nouvelle génération d'allocataires de l'APA. En sommant leurs évolutions nous obtenons une projection par GIR et type d'hébergement des effectifs de l'APA.

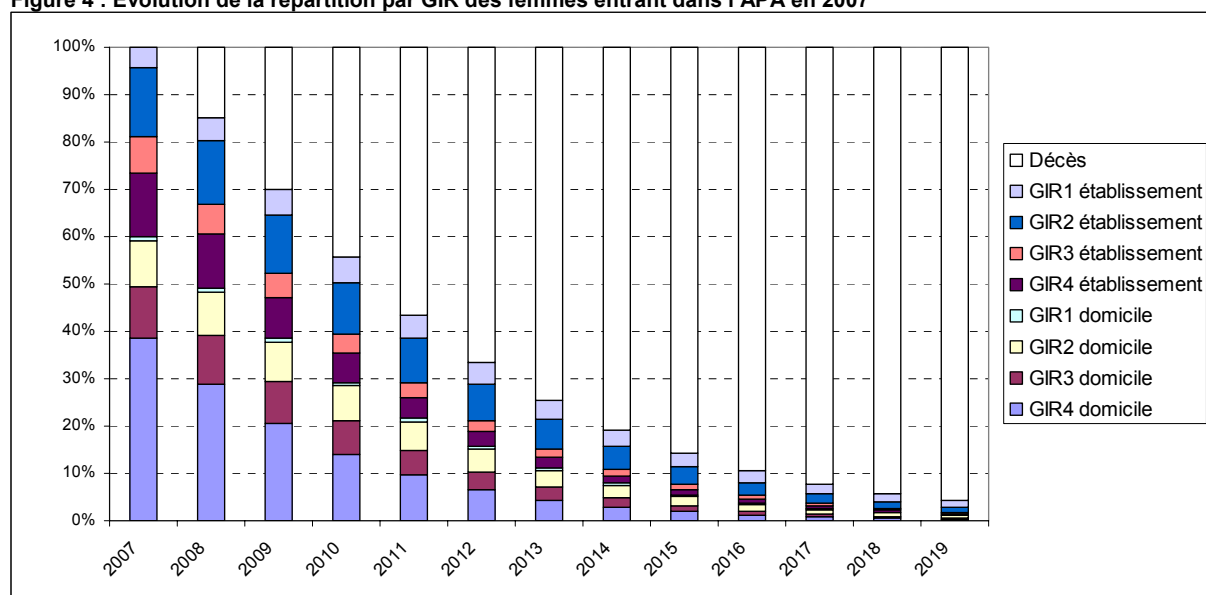
2.2.1. Les entrées dans l'APA

Pour étudier les entrées dans l'APA il faut tout d'abord résoudre le problème des personnes qui rentrent directement en établissement sous-dotations globale et qui ne sont pas observées : il faut repondérer nos données. La Drees a calculé que 40% des APA perçues en 2007 étaient des APA établissement. Nous faisons l'hypothèse que cette répartition entre domicile et établissement est la même pour les entrants. Ceci nous permet d'estimer des taux d'incidence par sexe et âge pour chaque état du système.

Nous pouvons aussi évaluer la répartition des entrées en fonction de l'état d'entrée et de l'âge. En se servant de cette distribution pour calculer une moyenne pondérée, nous estimons à 4,4 ans le temps

moyen de survie des femmes percevant l'APA en 2007 pour la première fois. Nous pouvons aussi estimer l'évolution de la répartition par état :

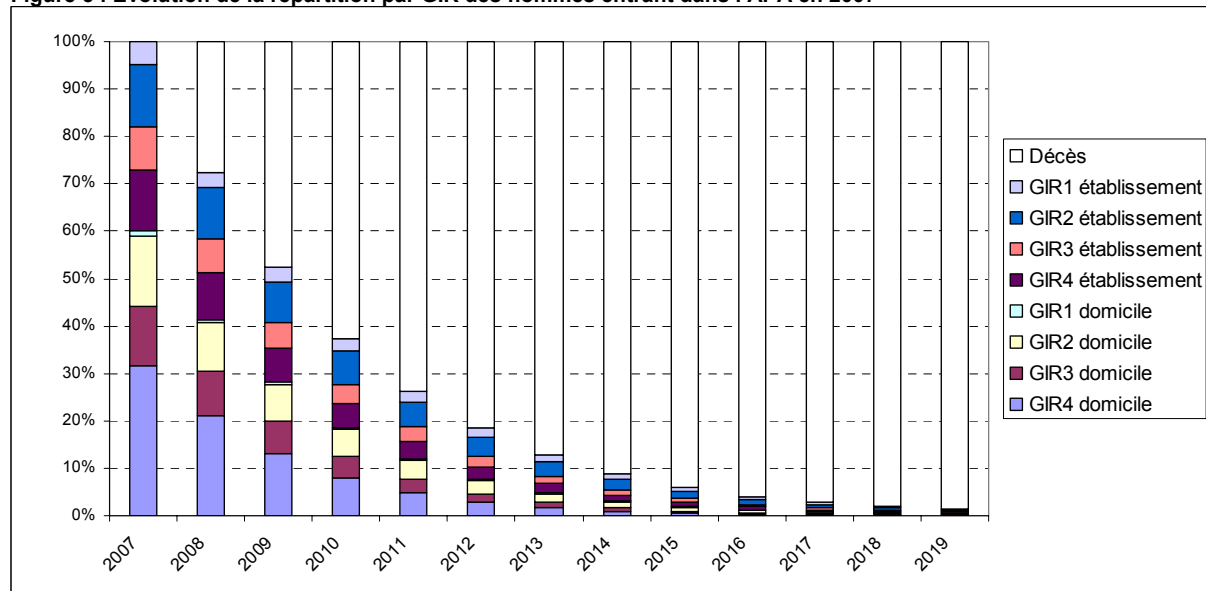
Figure 4 : Évolution de la répartition par GIR des femmes entrant dans l'APA en 2007



Source : Simulations de parcours à partir du modèle estimé automatiquement sur les données 2007

Au fur et à mesure des années passées dans l'APA, la part des personnes en établissement progresse et l'état GIR2 établissement devient plus fréquent que l'état GIR4 domicile. Pour les hommes on trouve une durée moyenne de 2,9 ans et on a la répartition par GIR suivante :

Figure 5 : Évolution de la répartition par GIR des hommes entrant dans l'APA en 2007



Source : Simulations de parcours à partir du modèle estimé automatiquement sur les données 2007

2.2.2. Survie des personnes percevant l'APA au 31 décembre 2007

Pour estimer l'évolution des effectifs des allocataires de l'APA, il faut évaluer le comportement du stock c'est-à-dire de l'ensemble des personnes qui perçoivent l'APA au 31 décembre 2007. Pour simuler un stock représentatif de l'APA en France (en particulier il faut éliminer le biais dû aux établissements sous dotation globale) nous avons calé la distribution de notre base de données (3 départements seulement) sur des répartitions nationales calculées par la Drees. Les critères utilisés

sont l'âge, le sexe et le type d'hébergement. Nous avons ensuite calculé le temps restant à percevoir l'APA. Pour les femmes, sous l'ensemble des hypothèses mentionnées au cours du document (et rappeler en synthèse à la fin) nous trouvons les résultats suivants :

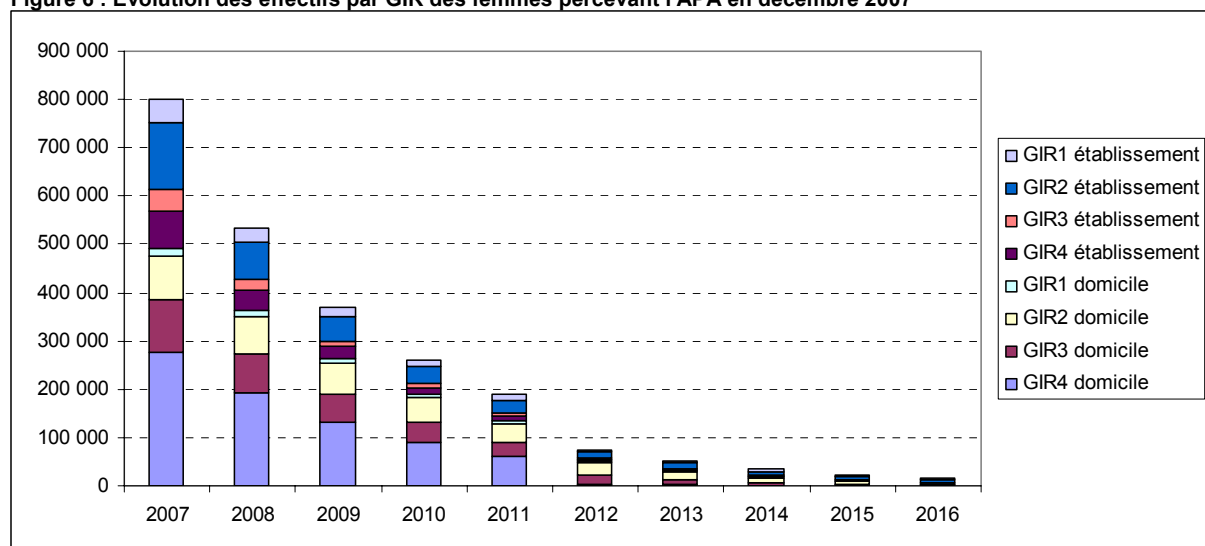
Tableau 9 : Temps moyen de survie pour les femmes percevant l'APA au 31 décembre 2007

GIR au 31 décembre 2007	Part des personnes percevant l'APA au 31 décembre 2007	Durée restante de perception de l'APA	Écart-type
Entrant GIR 4 domicile	33,3%	4,4	1,4
Autre GIR 4 domicile	1,2%	2,9	1,2
Entrant GIR 3 domicile	7,7%	4,1	1,6
Autre GIR 3 domicile	6,0%	2,9	1,0
Entrant GIR 2 domicile	6,1%	4,1	1,3
Autre GIR 2 domicile	5,2%	2,2	0,9
Entrant GIR 1 domicile	0,8%	5,7	3,0
Autre GIR 1 domicile	1,2%	1,7	0,9
Entrant GIR 4 établissement	4,9%	4,0	6,1
Autre GIR 4 établissement	4,7%	3,2	3,9
Entrant GIR 3 établissement	1,7%	3,5	4,0
Autre GIR 3 établissement	4,1%	2,9	3,5
Entrant GIR 2 établissement	5,5%	4,7	7,5
Autre GIR 2 établissement	11,7%	2,4	2,2
Entrant GIR 1 établissement	1,3%	12,2	13,0
Autre GIR 1 établissement	4,8%	1,8	1,3

Source : Simulations de parcours à partir du modèle estimé automatiquement sur les données 2007

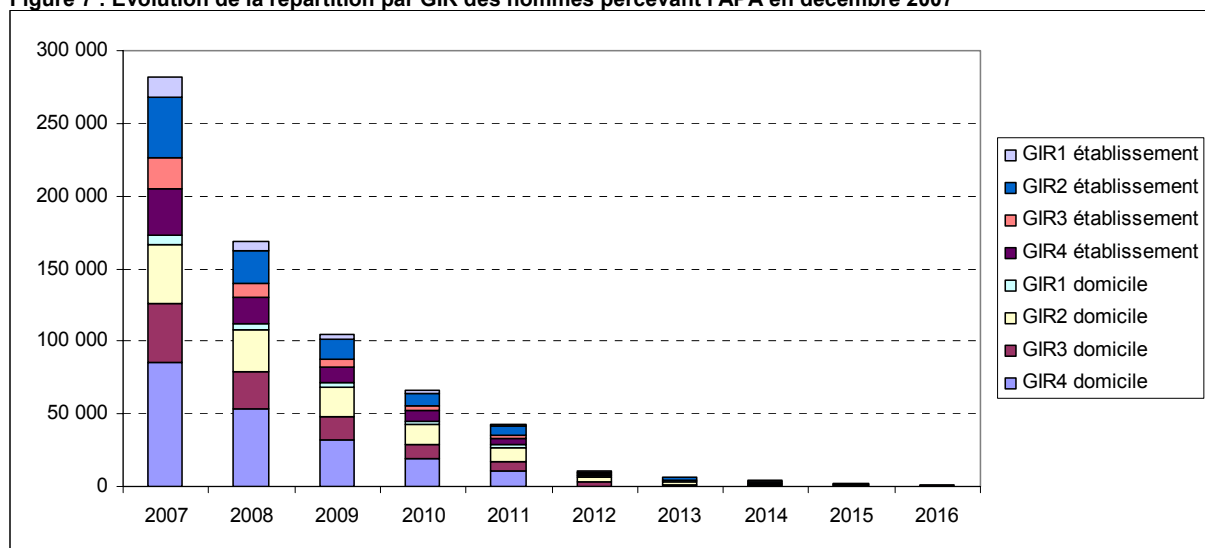
Ce temps dépend à la fois du GIR actuel, de l'âge mais aussi du temps déjà passé dans le dispositif de l'APA. En effet le modèle que l'on vient d'estimer rend possible une relation décroissante entre le temps que l'on a passé dans l'APA et celui qui reste à passer. Le modèle permet aussi d'estimer le parcours futur des personnes percevant l'APA à une date donnée. Les deux graphiques suivants illustrent l'évolution des effectifs par GIR des femmes et des hommes qui ont perçu l'APA en décembre 2007.

Figure 6 : Evolution des effectifs par GIR des femmes percevant l'APA en décembre 2007



Source : Simulations (modèle 2007), Base APA 2006-2007 et enquête annuelle sur l'APA Drees

Figure 7 : Évolution de la répartition par GIR des hommes percevant l'APA en décembre 2007

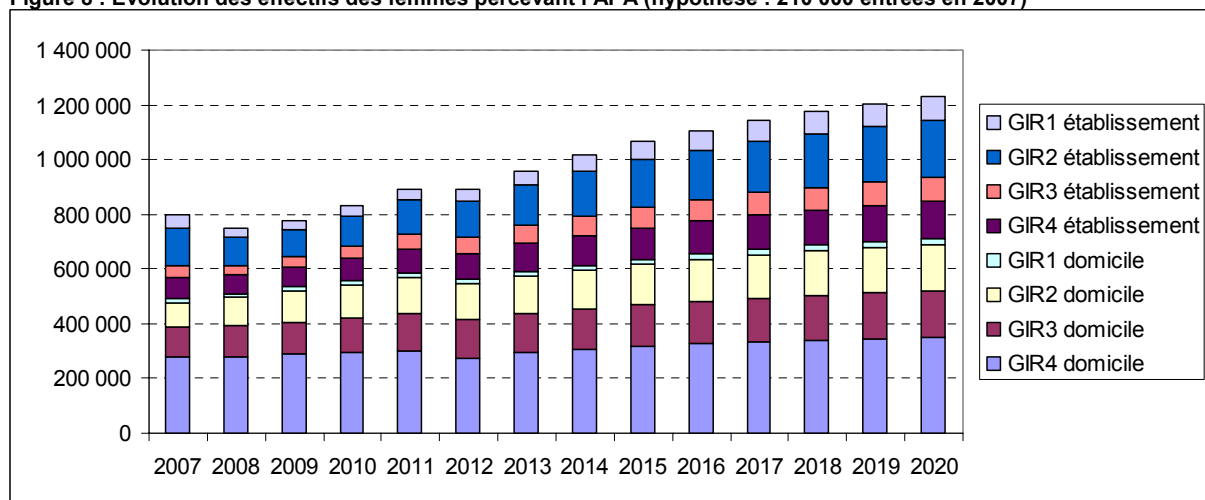


Source : Simulations (modèle 2007), Base APA 2006-2007 et enquête annuelle sur l'APA Drees

2.2.3. Projections : les problèmes soulevés

Lors du calage de la base des entrées 2007 dans l'APA, nous nous sommes servis d'une estimation du nombre total d'entrées calculée à partir de chiffres fournis par la Drees. Le flux d'entrants dans le dispositif a été estimé en considérant que le nombre d'entrants était proportionnel au stock pour une année et à lieu de vie donné. L'estimation obtenue est de 210 mille femmes et 90 mille hommes qui ont fait une demande d'allocation en 2007 qui a été acceptée. Ces chiffres influencent indirectement notre estimation de l'incidence des entrées dans l'APA (que l'on décompose bien sûr par GIR et type d'hébergement). Si l'on retient le chiffre de 210 mille entrées pour les femmes on obtient la projection démographique suivante :

Figure 8 : Evolution des effectifs des femmes percevant l'APA (hypothèse : 210 000 entrées en 2007)



Source : Simulations (modèle 2007), Base APA 2006-2007 et enquête annuelle sur l'APA Drees

Ce résultat pose problème car il montre une baisse des effectifs en 2008 ce qui paraît peu vraisemblable. Si l'on suppose le modèle juste, pour annuler la baisse simulée entre 2007 et 2008 il faudrait augmenter le nombre d'entrées (donc l'incidence) à 250 mille personnes en 2007 (soit une hausse de 19% par rapport à l'hypothèse de 210 000 entrées cette année là, alors que les effectifs globaux n'ont augmentés que de +3,8% entre septembre 2007 et septembre 2008).

3 Conclusion

Le travail que nous venons de présenter ne constitue qu'une première étape dans la mise en place d'un outil de prévision des effectifs de l'APA. Dans cette section nous allons rappeler les hypothèses techniques utilisées pour définir notre modèle et préciser son intérêt théorique. Nous rappelons ensuite les approximations faites pour exploiter les données partielles de trois départements. Ces données nous fournissent une estimation qui ne constitue qu'une illustration du potentiel de l'outil mis en place. Elle nous permet toutefois de révéler un certain nombre de problèmes restant à résoudre que nous énumérons dans la dernière partie.

3.1 Hypothèses utilisées pour construire le modèle

Le modèle que nous proposons est une simplification des mécanismes qui dirigent les trajectoires suivies par les allocataires de l'APA. La simplification s'applique tout d'abord à la dépendance temporelle : nous avons supposé que le parcours futur d'un allocataire ne dépend que de son état actuel et donc nous avons négligé les dépendances d'ordre supérieur. Cette hypothèse a été en partie relâchée en créant de nouveaux états. Chaque couple (niveau de GIR-type de logement) a été scindé en deux états : celui des entrants dans le système et celui des personnes déjà allocataires de l'APA. Nous avons ensuite simplifié les parcours dans l'APA. Le modèle se concentre sur les transitions qui nous ont paru significatives, en particulier nous avons négligé les passages d'établissement à domicile. Nous n'avons retenu qu'une variable explicative de la durée de perception de l'aide : l'âge d'entrée dans un état. Enfin nous avons estimé les durées passées dans chaque état avec un modèle paramétrique (Weibull ou exponentiel) ce qui ne permet de tester qu'un nombre très restreint de lois.

3.2 Intérêt théorique du modèle

La modélisation repose sur des hypothèses fortes qui restreignent la généralité de l'analyse, toutefois elle permet de résoudre plusieurs difficultés techniques. Elle permet en particulier, de répondre à la censure et à la troncature des données en analysant séparément le temps passé dans chaque état. Cette analyse séparée permet aussi de produire des résultats par type d'hébergement et par niveau de dépendance ce qui présente un intérêt car ces états sont associés à des niveaux de prestation différents. Enfin l'estimation d'un sous-modèle différent pour chaque état permet d'exploiter des données partielles comme celles des personnes passées en établissement sous dotation globale. Le modèle permet d'estimer une loi pour la dynamique du système d'allocation. En s'aidant de projections démographiques on est en théorie, capable d'évaluer les futurs effectifs de l'APA en donnant le détail par niveau de dépendance et type d'hébergement.

3.3 Approximations utilisées dans l'exploitation des données

Comme nous l'avons déjà rappelé les observations sont tronquées à gauche en 2002, pour éviter tout biais dû à des données de stock, nous avons utilisé une partie seulement de l'information disponible : pour chaque transition on n'a retenu que les données de la période d'observation. Pour les personnes passant en établissement sous dotation globale nous avons été obligé d'imputer l'état d'arrivée, en faisant l'hypothèse que le niveau de GIR en établissement était identique à celui atteint à domicile. Une autre série d'approximation sont nécessaires pour agréger les durées moyennes de perception de l'allocation ou pour faire des projections. Nous avons par exemple, calé les données des trois départements enquêtés sur la distribution nationale par GIR des entrants dans l'APA. Ce calage nous fournit une distribution par âge et par GIR des entrants ce qui permet d'obtenir des moyennes pondérées des durées de perception de l'allocation.

3.4 Limites du modèle et problèmes soulevés par les résultats

L'exploitation des données nous fournit une première estimation du modèle. Comme elle n'utilise qu'un échantillon très partiel de données elle ne constitue qu'une illustration du potentiel technique du modèle utilisé. Cette estimation nous permet cependant de révéler un certains nombres de problèmes qu'il faudra contrôler dans l'utilisation future des données :

- Disparités départementales : les résultats montrent qu'il existe des différences significatives entre départements pour la durée de perception de l'allocation. Il est donc nécessaire de prendre en compte ce critère pour mesurer la représentativité des données collectées.
- Durée de perception de l'allocation en GIR 1 établissement : les personnes en GIR 1 établissement constituent une population hétérogène où les durées de perceptions sont très variables et semblent dépendre de caractères inobservés qu'il faudrait préciser.
- Dépendance des estimations et du choix des observations : les estimations faites à partir de trois jeux de données (l'ensemble des observations de 2006 et 2007, les seules données de 2006 et la base des observations de 2007) montrent que les résultats évoluent au cours du temps. Même en prenant en compte la censure des données en 2002, nous obtenons des estimations plus importantes du temps de perception de l'APA à partir des données les plus récentes. Ceci relativise fortement le pouvoir prédictif du modèle mis en place.
- Utilisation d'un modèle simplifié pour les états d'entrée dans l'APA : nous rappelons que faute de temps nous présentons une estimation simplifiée du modèle. Elle est obtenue à partir d'une procédure d'estimation automatique appliquée aux données de l'année 2007. La procédure automatique utilise un même jeu de variables explicatives (des classes d'âges) fixé a priori. Il est possible d'effectuer une estimation plus rigoureuse en choisissant pour chaque état un modèle particulier : on peut choisir des classes d'âge significatives ainsi que le type de loi (soit une loi exponentielle, soit une loi de Weibull). Cependant du fait de la censure des données en 2002, les 8 états d'entrée dans l'APA présentent une difficulté particulière : nous n'avons pas réussi à estimer un modèle utilisant des paramètres de dispersion différents pour chaque sortie et nous avons calculé un modèle de Weibull avec un paramètre de dispersion unique. Le modèle avec des paramètres de dispersions différents reste toutefois possible pour les 8 autres états du système.
- Projections erronées sur les premières dates : l'étude conclue sur une évaluation des effectifs futurs qui paraît peu réaliste pour les premières années. Ce problème fondamental n'est pas résolu et les causes peuvent en être multiples. On peut par exemple penser à une mauvaise représentation des données nationales par les trois départements ou à une sous-évaluation des durées futures de perception de l'allocation (cette dernière hypothèse semble compatible avec la hausse constatée entre l'estimation fournie par les données 2006 et celle obtenue à partir des données 2007).

Bibliographie

- [1] ALLISON P. D., « Survival analysis using the SAS system: a practical guide » SAS Institute (Cary (NC)) 1995
- [2] CAMBOIS E., LIÈVRE A., « Risque de perte d'autonomie et chances de récupération chez les personnes âgées de 55 ans et plus » Études et Résultats n°349 novembre 2004 Drees
- [3] DORTET-BERNADET V., « Durée de perception de l'Allocation Personnalisée d'Autonomie : Mise au point d'un modèle de prévision », Rapport de Stage Ensae 2008
- [4] DUÉE M., REBILLARD C., « La dépendance des personnes âgées : une projection à long terme » Document de Travail de la Direction des Etudes et Synthèses Economiques de l'Insee (2004)
- [5] ESPAGNOL P., LO S-H., DEBOUT C., « L'allocation personnalisée d'autonomie et la prestation de compensation du handicap au 30 juin 2008 », Etudes et Résultats n° 666 octobre 2008
- [6] LANCASTER T., « The econometric analysis of transition data », Cambridge University Press (Cambridge (GB)) 1990
- [7] PEROU O., « L'assurance dépendance » Mémoire diplôme d'économiste Cnam en actuariat
- [8] PLANCHET F., THÉRON P., « Modèles de durée : applications actuarielles » Economica (Paris) 2006