

PRISME
**Le modèle de microsimulation dynamique de la CNAV,
régime français de pension de retraite**

Christophe Albert¹ (CNAV)
Caroline Berteau-Rapin² (CNAV)
Alessandra Di Porto³ (CNAV)

¹ Christophe.albert@cnav.fr

² Caroline.berteau-rapin@cnav.fr

³ Alessandra.di-porto@cnav.fr

Introduction la genèse du modèle de microsimulation <i>PRISME</i>	3
A- Echantillon, complétions, architecture	5
1 Présentation de l'échantillon et complétions.....	5
1.1 Présentation de l'échantillon	5
1.2 Les complétions.....	6
2 Organisation de la projection	9
B- Modélisation (blocs, sous-blocs).....	11
1 L'ajout de nouveaux cotisants à la base	11
2 La modélisation des évènements démographiques	12
2.1 Les naissances	12
2.2 Les décès	13
3 La modélisation de l'activité et des salaires.....	14
3.1 Modélisation de l'activité.....	14
3.2 Modélisation des salaires	16
4 La modélisation des départs en retraite	17
4.1 Les pensions des non-réclamés	18
4.2 Les départs avant 60 ans (ou départs en retraite-anticipée)	18
4.3 Les départs à partir de 60 ans	19
5 La modélisation de la réversion	22
C- Déroulement type et résultats	25
1 Déroulement type	25
2 Résultats et simulations	27
2.1 Résultats – scénario central.....	27
2.2 Simulation du décalage de l'âge légal de 60 à 61 ans.....	31
Conclusion – Avenir de <i>PRISME</i>	34
Annexe	35
Glossaire.....	38
Bibliographie.....	42

*"Pour l'enfant, amoureux de cartes et d'estampes,
L'univers est égal à son vaste appétit.
Ah ! que le monde est grand à la clarté des lampes !
Aux yeux du souvenir que le monde est petit !"*
(Charles Baudelaire)

Introduction

la genèse du modèle de microsimulation *PRISME*

En préambule à la présentation du modèle de projection *PRISME* créé par notre équipe, il nous semble nécessaire de rappeler le cadre institutionnel dans lequel il a été développé, et ce qui a motivé sa création en 2004.

Notre employeur, le régime général (*CNAV*), est le principal régime de base de retraite en France. Il couvre les salariés du secteur privé français, et gère en répartition actuellement environ 12 millions de retraités et 17 millions de cotisants (71% des actifs). Seulement 3% des personnes connues du système de protection sociale français ne passent pas par le régime général au cours de leur vie. Celui-ci fait partie d'un ensemble de régimes de base, chacun regroupant les assurés selon un découpage professionnel que l'on peut résumer ainsi : salariés, artisans-commerçants, professions libérales, secteur public et régimes spéciaux. Le calcul de la pension est détaillé dans l'encadré 1.

Les assurés actifs cotisent aujourd'hui sur une base de 16,65% du salaire brut, dont 1,7% pour la partie supérieure au plafond de la sécurité sociale fixé aujourd'hui à 2 900€ par mois. La pension moyenne des nouveaux retraités de l'année s'établit à environ 600€, complétés par la pension de régimes complémentaires et éventuellement des autres régimes de base. La pension mensuelle moyenne tous régimes pour l'ensemble des retraités des régimes français s'établit désormais à plus de 1100 €.

La stagnation attendue des cotisants et l'envolée prévue du nombre de retraités ont contraint les pouvoirs publics à adapter les dispositifs de retraite afin de réduire, sinon d'éviter, la dégradation de l'équilibre des régimes. Dans cet esprit, une réforme des retraites a été engagée en 2003. Cette réforme très technique a rendu obsolètes les modélisations basées sur des individus moyens utilisées jusqu'alors.

Un modèle de microsimulation dynamique nommé *PRISME* (Projection des Retraites, Simulations, Modélisation et Evaluations) a par conséquent été créé à la *CNAV*, afin d'évaluer les conséquences globales et individuelles des changements de législation passés (et futurs). Micro-simulation parce qu'il s'appuie sur plus de 4 millions d'individus, dynamique en raison du nombre de périodes trimestrielles projetées (horizon 2050). Par ailleurs, ce modèle a été développé sous le langage SAS.

Encadré 1. Calcul de la pension de retraite servie par la CNAV

La pension de droit propre versée par le régime général est calculée comme le produit de trois paramètres : le salaire annuel moyen, le taux de liquidation de la retraite et un ratio représentant la présence au régime général dans l'ensemble de la carrière.

$$Pension = \text{Salaire annuel moyen} * \text{Taux} * \frac{\text{Durée assurance au RG}}{\text{Durée pour l'obtention du taux plein}} \quad (+\text{différentiel lié à minima éventuel})$$

En France, l'âge minimal de départ en retraite pour les assurés relevant du Régime général est de 60 ans. Liquider à 65 ans, ou en invalidité, ou encore après la durée statutaire tous régimes (de 40 à 41 années selon la génération) assure l'obtention du taux plein (50%) de liquidation et potentiellement d'un minimum de pension (pour 40% des assurés aujourd'hui). Décote ou surcote s'appliquent sur ce taux de liquidation en cas de carrière insuffisante ou de poursuite d'activité après l'obtention du taux plein (+/- 5% par an).

Depuis 2004, les assurés peuvent obtenir une retraite à partir de 56 ans à l'issue d'une longue carrière commencée très jeune.

Le salaire annuel moyen est calculé comme une moyenne des 25 meilleurs salaires, plafonnés à un niveau équivalent au salaire moyen, et revalorisés en fonction de l'inflation et non de la croissance du salaire moyen. Les pensions sont indexées suivent les prix une fois liquidées.

La CNAV verse également une pension aux veuves et veufs des assurés du régime, basée sur 54% de la pension du décédé, sous condition de ressources.

Dans une première partie, nous présenterons succinctement les méthodes d'échantillonnage et de complétion des données nécessaires à l'exercice de projection, ainsi que l'architecture retenue pour le modèle.

La deuxième partie présentera la modélisation des différentes étapes de *PRISME* (démographie, activité, salaires, départs en retraite, réversion), avec un zoom sur un module-clé, celui décidant de la date de liquidation de la retraite de chaque assuré encore actif en début de projection.

Une troisième et dernière partie présentera à la fois le déroulement-type d'une projection réalisée sous SAS et des exemples d'utilisation de la microsimulation dynamique.

A- Echantillon, complétions, architecture

1 Présentation de l'échantillon et complétions

1.1 Présentation de l'échantillon

Les données initiales sont des données de gestion alimentées en temps réel (identité, carrière, et dossiers des assurés s'ils sont retraités) ou des données statistiques relevées chaque trimestre. La table statistique regroupe un plus grand nombre d'informations sur les 12 millions de retraités que les seuls dossiers (types de pension, montants, date de perception, durées d'assurance...). Elle est restreinte aux seuls retraités du régime, tandis que les autres données concernent les assurés nés entre 1900 et 2008 et connus de la sphère de la protection sociale en France (assurances vieillesse, maladie, famille...).

PRISME s'appuie sur un échantillon au 20^{ème} actualisé tous les deux ans, ou plus précisément une sélection de 5 clés de contrôle prises au hasard parmi 97. Ces clés sont codées sur les deux chiffres qui suivent le *NIR*⁴. Ce mode de tirage de l'échantillon permet de conserver le panel constitué entre deux extractions, enrichi des nouveaux assurés correspondant au critère de sélection (naissances et immigrants). Les données portant sur la retraite sont actualisées tous les trimestres (évolutions de pension, décès...), pour une meilleure précision des prévisions de court terme.

La table initiale est créée selon une structure comportant autant d'observations que d'individus. Elle comporte plus de 800 variables. Chaque individu est repéré par un numéro d'ordre (le *NIR* est masqué pour des raisons de protection de l'anonymat), et l'ensemble de ses caractéristiques est enregistré dans des variables relatives à la démographie, la carrière et la retraite.

La partie démographie recense les informations générales sur l'assuré (sexe, année et trimestre de naissance, année et trimestre de décès éventuellement), ainsi que les données sur la descendance (nombre d'enfants, chronologie des naissances avec les trimestres et les années de naissance jusqu'au 5^{ème} enfant) et son statut d'immigré éventuel (pays de naissance, année et trimestre d'entrée sur le territoire français).

La partie carrière comprend les variables relatives à la carrière de l'individu : des variables relatives au début de carrière (âge de début d'activité, âge de fin d'études), des variables précisant le type de report et le nombre de trimestres associés pour chaque année et trimestre définissant la carrière, ainsi que les montants bruts de salaires annuels.

La partie retraite inclut les données nécessaires à l'estimation de la date de départ en retraite et les éléments utiles au calcul de la pension, notamment :

- le nombre de trimestres cumulé tout au long de la carrière au Régime Général et dans les autres régimes, avec une distinction entre durée validée et durée cotisée imposée par la dernière réforme,

⁴ Numéro d'Inscription au Répertoire (Répertoire national d'Identification des Personnes Physiques) : identifiant composé de 13 chiffres, appelé plus couramment "numéro de sécurité sociale", suivi d'une clé de vérification établie sur 2 chiffres.

- le Salaire Annuel Moyen (SAM),
- le taux de liquidation,
- le nombre de trimestres de surcote, etc.

Pour les assurés déjà en retraite, l'année et le trimestre de départ en retraite sont bien entendu enregistrés, ainsi que les montants de pension (avant et après minima de pension) et les aides sociales versées par le régime appelées "avantages complémentaires".

Parmi les informations listées, certaines ne figurent ni dans les données de gestion disponibles ni dans la base statistique des prestataires. Les raisons sont multiples :

- imperfection des données, problèmes de qualité ou d'alimentation (identification, données sur les revenus financiers...)
- renseignement des données uniquement à partir des 55 ans de l'assuré (carrière sous certains autres régimes professionnels, données individuelles sur la composition de la famille...)
- informations inutiles au système de protection sociale mais indispensables à la microsimulation (âge de fin d'études)

L'exercice de projection ne peut débuter sans une étape dite de "complétion" des données qui aboutit à une table initiale prête à l'emploi. Cette étape est décrite dans la partie suivante.

1.2 Les complétions

La complétion est à différencier de la création d'observations supplémentaires liées aux naissances et immigrations futures. Elle vise à créer des variables ou à les compléter, et non à ajouter des lignes à la table comme le fait la projection (*cf.* point 2).

Complétion des décès

La CNAV connaît a priori l'immatriculation (à la naissance pour les nés en France, à l'accès à la sphère de la protection sociale pour ceux nés hors de France) et l'occurrence du décès des résidents.

Cependant, une partie des décès n'est pas enregistrée,

- soit parce qu'il s'agit d'individus nés et morts il y a très longtemps,
- soit parce qu'il s'agit d'immigrés qui sont sortis du territoire français, et pour lesquels le décès n'est pas toujours connu.

Pour ces raisons, nous procédons à une complétion des décès pour les individus nés en France avant 1946 et pour ceux nés à l'étranger avant 1980. Nous calculons d'abord l'âge atteint au cours de l'année la plus récente pour laquelle une information est connue (emploi, chômage, maladie, retraite, ...).

Nous calculons ensuite, en fonction de la génération et du genre, les probabilités de décès à chaque âge⁵, sachant que l'individu est vivant à l'âge décrit précédemment. Ces probabilités sont des probabilités « ajustées », pour prendre en compte les décès déjà connus et les individus pour lesquels nous sommes sûrs du fait qu'ils sont encore en vie.

⁵ A partir des tables de l'INSEE (Institut National de la Statistique et des Etudes Economiques)

Nous déterminons enfin, par comparaison avec un aléa, l'éventuel décès d'un de ces assurés au titre de la complétion des données. Ce décès attribué est différent de l'événement décès simulé en projection (*cf.* partie architecture).

Complétion des âges de fin d'études

L'âge de fin d'études est une information déterminante pour de nombreux événements touchant l'assuré en projection. Il impacte notamment l'arrivée des enfants, la croissance salariale des assurés, les transitions autour de l'emploi et la mortalité. Cette variable, absente des fichiers administratifs de la Cnav, est complétée à l'aide des salaires contenus dans les fichiers.

Schématiquement, une forte hausse de salaire avant 30 ans est considérée comme un signal de fin d'études l'année précédant cet accroissement. Le taux de croissance du salaire a été calculé afin que la distribution des âges de fin d'études ainsi déterminés corresponde à celle déclarée par les salariés du secteur privé dans une enquête *INSEE* nommée "enquête emploi".

Pour les assurés dont les reports de salaires annuels ne connaissent pas une hausse jugée suffisante pour y reconnaître un signal utilisable, ou pour ceux qui n'ont pas de report de salaire avant 30 ans, des âges de fin d'études leur sont affectés afin que la répartition globale respecte celle de l'ensemble des déclarants de l'enquête-emploi.

Complétion des naissances

Les caractéristiques familiales, en particulier le nombre et l'âge des enfants des femmes, font partie des éléments indispensables pour la détermination des droits à la retraite future. En effet, outre le fait que ces éléments aient une influence sur le parcours professionnel (principalement des femmes), ils interviennent également dans la détermination des avantages familiaux versés par le régime général.

Au moment de la constitution de la base, la descendance atteinte par chaque assuré l'année de début de la projection est intégrée dans le modèle de projection, en fonction de ses caractéristiques personnelles.

La descendance des femmes est estimée à partir des caractéristiques individuelles (génération, pays de naissance et âge de fin études) et des données de carrière (validation de reports maladie ou maternité ou validation de trimestres au titre des congés parentaux). Ces dernières permettent de déterminer les années « éligibles », celles pour lesquelles les informations de carrières sont compatibles avec une éventuelle naissance.

Nous déterminons donc, pour chaque année éligible, la probabilité d'avoir un enfant avec les équations logistiques calculées par l'Institut National de Démographie (*INED*), sur la base d'une enquête nationale relative à l'histoire familiale (*'Enquête Famille'* de 1999).

La comparaison de cette probabilité à un aléa déterminera une éventuelle naissance.

Nous attribuons une descendance aussi aux hommes, ces derniers pouvant également bénéficier d'un avantage familial lié au nombre d'enfants. La méthode d'attribution des enfants aux hommes est plus simple : seul le nombre d'enfants est précisé, non leur année d'occurrence. La probabilité d'avoir des enfants est calculée en respectant les informations relative à une éventuelle liquidation de la retraite (présence ou pas d'un avantage familial qui permet de savoir si l'assuré a eu au moins trois enfants), et en respectant la répartition

observée dans l'enquête famille exploitée par l'*INED*. Elle prend en compte la génération, le pays de naissance et l'âge de fin études du père.

Complétion des carrières

Certaines périodes d'activité relevant d'autres régimes sont absents des fichiers de la Cnav. Il s'agit des périodes militaires et des périodes validées dans un autre régime.

La complétion des carrières des individus se décompose en trois parties :

- la complétion des périodes militaires,
- la complétion des reports validés dans les autres régimes par les individus déjà passés au régime général,
- la complétion des carrières des individus qui ne sont encore pas passés au régime général.

La complétion des périodes militaires s'effectue sur les hommes âgés de 18 à 30 ans à l'aide des données de carrière validées par les individus ayant déjà liquidé ainsi que des effectifs des militaires du contingent du Ministère de la Défense. A partir de ces effectifs d'individus présents une année donnée sous les drapeaux, en fonction de la répartition par âge, et de la durée du service, on calcule des effectifs de nouvelles incorporations annuelles.

Pour une année d'incorporation donnée et une génération donnée, les individus sélectionnés auxquels sont attribuées des périodes assimilées militaire sont ceux présentant une durée d'absence de report correspondant à la durée courante du service militaire.

La complétion des reports validés dans les autres régimes par les individus déjà passés au régime général s'effectue à l'aide de l'Echantillon Inter-régime de Cotisants de la *DREES*⁶ et d'équations logistiques.

Cette complétion s'effectue en trois étapes : sélection des individus polypensionnés, puis choix du type de régimes à l'origine de ces reports autres régimes, et enfin imputation d'une durée autres régimes validés. Chacune de ces étapes comporte trois phases : une phase d'estimation des coefficients des équations à partir d'individus dont la carrière est intégralement connue, une phase d'estimation de la variable réponse pour les individus à compléter et une phase de sélection des individus en fonction de la valeur de leur variable réponse et de l'effectif sur lequel nous devons nous caler.

La complétion des carrières des individus qui ne sont pas passés au régime général s'effectue en sélectionnant les individus n'ayant aucun type de reports renseigné dans leur carrière. Les données de l'EIC nous permettent de calculer le rapport entre le nombre d'individus ayant validé des reports dans un régime aligné (et non aligné) sans valider de trimestre au régime général et le nombre d'individus ayant validé au moins un trimestre au régime général. En appliquant ce rapport à l'effectif des individus ayant déjà validé un trimestre au régime général dans notre échantillon, on obtient le nombre d'individus n'ayant encore rien validé au régime général que l'on doit compléter en reports venant d'autres régimes.

Enfin, des calages transversaux sont réalisés afin que le nombre d'individus qui effectuent une année donnée une activité dans un autre régime soit conforme à celui obtenu par des sources externes.

⁶ Direction de la Recherche, des Etudes, de l'Evaluation et des Statistiques, Ministère de la Santé et des Sports

2 Organisation de la projection

Une fois la table initiale constituée, le travail de préparation n'est pas achevé. Il est nécessaire d'entrer tous les paramètres propres à la projection, à savoir :

- Les hypothèses démographiques, notamment les tables de mortalité adoptées (parmi plusieurs possibles), le taux de fécondité à l'avenir, le nombre d'immigrations nettes sur le territoire français jusque 2050. Les lignes créées pour les naissances et immigrations futures sont ajoutées à la table initiale.
- les hypothèses macroéconomiques sur le même horizon temporel, en particulier la croissance supposée du salaire moyen et le taux de chômage de la population. Au cours de la projection, nous calculons chaque année le nombre de chômeurs et vérifions que le taux de chômage ainsi établi correspond au taux souhaité. Nous procédons de manière identique pour le nombre de cotisants de notre régime, une fois déterminé le nombre d'actifs occupés. Des calages sont éventuellement incorporés, qui demandent une vérification du respect de la croissance du salaire moyen.
- Les hypothèses correspondant au type de législation retenu (avant 1993, avant 2003, depuis 2003 schématiquement et en rapport avec les deux grandes réformes qu'a connues le régime général). L'indexation des minima de pensions est également à préciser, parmi d'autres montants nécessaires au calcul des pensions. Ainsi, un grand nombre de tables sont entrées en "mémoire" dans SAS, correspondant aux taux de surcote, de décote, au nombre de trimestres à atteindre pour obtenir le taux plein (par génération). La partie "déroulement concret d'une projection" précise comment ceci est réalisé.
- Enfin, tous les paramètres des équations logistiques ou des matrices de transition sont mis sous format ou entrés sous forme de macro-variable afin de calculer les probabilités d'occurrence de tous les événements que peuvent connaître les assurés. Cela sous-entend que l'ensemble des travaux économétriques nécessaires ont été réalisés avant projection.

B- Modélisation (blocs, sous-blocs)

Cette partie a pour but la présentation de la modélisation des différentes étapes de *PRISME*. Dans une première sous-partie nous présentons l'ajout des nouveaux cotisants à la base de projection, puis dans une deuxième, la modélisation des événements démographiques. Une troisième sous-partie sera consacrée à la modélisation de l'activité et des salaires des individus présents dans notre base. Enfin, dans une quatrième et cinquième sous-partie seront présentées respectivement les modélisations des départs à la retraite et de la réversion. Un zoom est effectué sur le module présenté dans la quatrième partie, module-clé décidant de la date de liquidation de la retraite de chaque assuré encore actif en début de projection.

1 L'ajout de nouveaux cotisants à la base

Lors du lancement du programme de projection, la première étape qui est exécutée consiste à créer les individus qui apparaîtront après le début de la projection, qu'ils soient nés en France ou à l'étranger.

Future births in France

Le nombre de nouveaux assurés à intégrer par génération et sexe est calculé à partir des prévisions des naissances jusqu'en 2050 faites par l'Institut national des statistiques (INSEE). Les naissances ainsi obtenues sont réparties entre les deux sexes selon la moyenne des rapports de masculinité à la naissance constatés (51,19% de naissances de sexe masculin, contre 48,81% de naissances de sexe féminin).

Born abroad future contributors

Des « lignes » correspondant aux individus nés à l'étranger et arrivant sur le territoire français sur la période de projection sont ajoutées dans la base, afin d'intégrer cette population à celle des cotisants.

Pour estimer le nombre d'immigrés qui arrivent en France nous nous appuyons sur le nombre d'assurés nés hors France ayant validé un premier report, quel que soit le type (report pour emploi, chômage ou autres périodes assimilées), sur les dix années précédant le début de la projection.

Le flux des premiers reports est très proche du flux légal estimé par l'Ined (Institut National d'Etudes Démographiques), bien que les deux populations prises en compte ne soient pas les mêmes : il s'agit de l'immigration avec une durée de séjour supérieure à un an pour l'Ined, et de l'immigration pour travail (même saisonnier) ou des reports dus aux périodes assimilées pour la Cnav.

Le flux d'immigrés est estimé à 200 000 entrées par an. Ces entrées sont réparties de manière égale entre les deux sexes.

La distribution par âge du flux d'immigrés adoptée est celle qui correspond à la distribution issue du Recensement général de la population de 1999 pour les années 1997 et 1998 relative aux personnes immigrées en France.

Nous attribuons à ces immigrés un âge de fin d'études et une fécondité selon les mêmes procédures utilisées pour les nés en France. En ce qui concerne le nombre de trimestres validés avant l'émigration, nous attribuons un nombre de trimestres qui est fonction des pourcentages constatés sur le passé pour les autres assurés nés à l'étranger.

2 La modélisation des événements démographiques

Le modèle de projection intègre la modélisation de deux événements démographiques : les naissances et les décès que nous allons développer dans les sous-parties suivantes.

2.1 Les naissances

Le nombre d'enfants et le calendrier des naissances sont attribués aux femmes nées en France d'âge compris entre 15 ans et 49 ans, selon des équations logistiques, en fonction du temps passé depuis la fin des études.

L'arrivée du premier enfant est modélisée différemment de celle des éventuels frères et sœurs.

$$P(1^{\text{ère}} \text{ naissance}) = \frac{1}{1 + \exp[-(\alpha + \beta \text{ âge} + \gamma \text{ âge}^2 + \delta \text{ âge}^3 + \eta \text{ durée fin études} + \vartheta \text{ durée fin études}^2)]}$$

$$P(\text{naissance rang} > 1) = \frac{1}{1 + \exp \left[\begin{array}{l} -(\alpha + \beta \text{ âge} + \gamma \text{ âge}^2 + \delta \text{ âge}^3 + \eta \text{ âge benj} \\ + \vartheta \text{ âge benj}^2 + \omega \text{ âge benj}^3 + \phi \text{ âge benj}^4) \end{array} \right]}$$

Ainsi, par exemple, la probabilité d'une première naissance est donnée par les paramètres suivants :

Tableau 2.1.1 - Probabilité annuelle d'une première naissance

Paramètres	Fin d'études < 18	Fin d'études > 17
Constante (α)	-14,7458	-57,9839
Age de la femme (β)	1,2729	5,069
(Age de la femme) ² (γ)	-0,045	-0,1518
(Age de la femme) ³ (δ)	0,00051	0,00148
Durée depuis fin études (η)	0,2972	0,2185
(Durée depuis fin études) ² (θ)	-0,0145	-0,0122

Ces paramètres ont été calculés conjointement avec l'Institut national d'études démographiques (INED), à partir de l'Enquête famille de 1999.

Plusieurs séries de paramètres ont été calculées, selon différents indices conjoncturels de fécondité (ICF). Nous appliquons actuellement un ICF de 1,9 enfants par femme.

Pour chaque femme et pour chaque année de vie féconde, nous calculons la probabilité qu'elle ait un enfant, en fonction des valeurs des variables qui sont appliquées dans l'équation logistique. Si l'aléa tiré est inférieur à la probabilité ainsi calculée, nous augmentons la descendance de l'assurée d'un enfant et nous attribuons une date de naissance de l'enfant.

Ces équations sont également appliquées aux femmes nées à l'étranger, pour les années antérieures à l'immigration sur le territoire français (dans ce cas, nous considérons que les naissances peuvent se produire jusqu'à un an et demi avant l'entrée en France) et pour celles postérieures à l'entrée en France.

Afin d'obtenir une fécondité légèrement plus élevée pour les femmes nées à l'étranger par rapport aux femmes nées en France, et en raison de la modélisation des naissances retenue dans le modèle Prisme, les immigrées sont supposées être peu qualifiées quelque soit l'âge de fin d'études.

En ce qui concerne les hommes nés en France, aucun enfant ne leur est attribué : en effet, les hommes nés avant le début de la projection ont un nombre d'enfants qui est renseigné lors de la complétion des naissances et ceux qui vont naître après le début de la projection ne seront pas encore prestataires à la fin de la projection.

Nous attribuons aux hommes nés à l'étranger un nombre aléatoire d'enfants, en respectant les répartitions des hommes selon leur génération et le nombre d'enfants qu'ils ont eu (répartitions données par l'Ined).

2.2 Les décès

Les prestataires titulaires d'une pension pour inaptitude ou les ex-invalides se caractérisent par un niveau de la mortalité plus élevé que les assurés titulaires d'une pension « normale ». L'écart entre l'espérance de vie à 60 ans des titulaires d'une pension normale et celle des retraités au titre de l'invalidité ou de l'inaptitude est au minimum de 5 années. A l'opposé, l'espérance de vie à 60 ans des titulaires d'une pension normale est plus élevée que celle de la population française globale.

Au niveau global, l'espérance de vie à 60 ans toutes catégories de pensions confondues est très proche de celle calculée au niveau national par l'Insee.

Afin de mieux prendre en compte la mortalité des prestataires, nous calculons, à partir de nos bases des données, des séries des quotients de mortalité par sexe, âge et année, déclinées selon le type de pension (normale, inapte ou ex-invalidé).

Les séries des quotients de mortalité sont calculées jusqu'en 2005 et ensuite prolongées jusqu'en 2050, en suivant l'évolution de la mortalité de la population française.

Nous calculons alors, à partir de ces séries de quotients de mortalité, la probabilité que chaque prestataire de droit propre a de décéder au cours d'une année, en fonction de son sexe, de son âge et du type de pension dont il bénéficie.

Les assurés qui ne sont pas encore prestataires sont soumis, quant à eux, à des quotients de mortalité calculés à partir des projections de population et de décès à 2050 réalisées par l'INSEE.

Une fois les quotients de mortalité calculés, les prestataires sont soumis à la probabilité de décéder pour chaque année de la projection. Nous comparons ainsi un aléa à la probabilité

« théorique » de décéder, relative à leurs caractéristiques (sexe, âge, catégorie) : si l'aléa tiré est inférieur à la probabilité de décéder, nous attribuons une date de décès.

Une fois décédé, l'individu ne rentre plus dans les boucles de programme suivantes, excepté pour la réversion.

Une trimestrialisation des décès est prise en compte, afin de reproduire la saisonnalité des décès constatée et d'améliorer les prévisions de court terme (22% de décès en plus en janvier par rapport à juin).

3 La modélisation de l'activité et des salaires

Trimestre par trimestre, nous dessinons l'avenir des assurés, tant en termes de type d'activité, que de niveau de salaire en cas d'activité dans notre régime.

3.1 Modélisation de l'activité

Le module de transition, permet de modéliser des statuts d'activité de chaque individu pour chaque année et chaque trimestre projeté. Il permet de simuler ainsi le partage de la population entre actifs et inactifs, et au sein des actifs, entre salariés du Régime Général, salariés d'autres régimes (alignés et non alignés), chômeurs, salariés en arrêt maladie. Ainsi, à chaque pas de la projection, l'une des situations suivantes est affectée à chaque individu en âge de travailler (14-69 ans) :

- en activité au Régime Général,
- en activité dans d'autres régimes alignés (salariés de l'agriculture, artisans, commerçants),
- en activité dans d'autres régimes non alignés,
- au chômage,
- en maladie avec indemnités journalières,
- en invalidité ou accident du travail,
- dans une autre situation ne permettant pas de valider de trimestre.

Actuellement une matrice de transitions est utilisée afin de modéliser les statuts d'activité des individus.

Cette matrice, basée sur les reports observés sur le passé (actuellement sur les années 2001-2003) donne par genre, tranche d'âge, nature du report affecté le trimestre précédent et pays de naissance la probabilité de passer trimestriellement d'un état vers un autre. Les bornes des tranches d'âge varient en fonction de l'âge de fin d'études afin de prendre en compte l'allongement de la durée des études constaté ces dernières années.

A titre d'exemple, un homme né en France, ayant terminé ses études quelques années auparavant, sortant d'un report d'activité au régime général au trimestre précédent a :

- 93% de chance d'avoir un nouveau salaire sur le trimestre courant,
- 3% de chance de ne valider aucun report,
- 3% de chance de valider une période de chômage indemnisé,
- moins de 1% de chance d'avoir une période assimilée maladie,
- environ 1 chance sur 1000 d'aller dans un autre régime aligné ou non aligné,
- une probabilité quasi nulle de devenir invalide.

En projection, ces probabilités de transitions ne sont pas rigoureusement appliquées. Les organismes à l'origine de la demande de projection fixent un certain nombre de contraintes concernant le sentier macroéconomique à suivre.

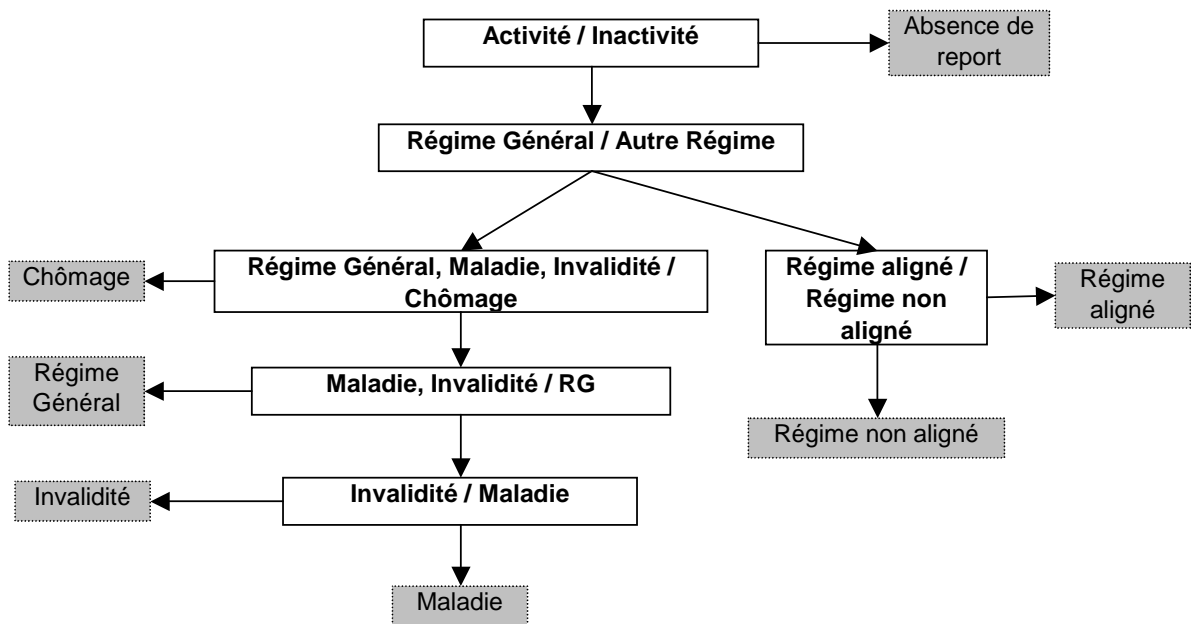
Cela se traduit par cinq types de calages :

- personnes au chômage,
- cotisants du Régime général,
- cotisants dans un régime aligné,
- cotisants dans un régime non aligné,
- personnes en invalidité.

Cependant, dans un souci d'amélioration continue du modèle, et afin de retranscrire de manière plus satisfaisante les trajectoires individuelles, des travaux sont en cours pour introduire un module de transitions construit sur la base d'un enchaînement d'équations logistiques permettant d'intégrer plus précisément les caractéristiques individuelles et la trajectoire passée de l'assuré.

Cet enchaînement de logistiques dichotomiques se décline également par genre, distance à l'âge de fin d'études et pays de naissance, soit au total 96 équations.

L'enchaînement des équations logistiques est schématisé ci-dessous :



Note de lecture : Les cases blanches correspondent aux différentes situations testées. Les cases grises sont les reports affectés comme résultat de chaque test.

A chaque transition, deux situations exclusives sont testées : la probabilité d'être ou non dans tel ou tel état sur le trimestre en cours est comparée à un aléa qui permet soit d'imputer la situation de référence, soit d'enchaîner avec le calcul de la probabilité de transition de la situation suivante.

Les équations de transition sont déterminées à partir des données contenues dans la table initiale du modèle de projection, sur la période 2001-2003.

Les variables explicatives retenues varient selon la transition considérée, les groupes d'âge, le genre, le pays de naissance et la significativité de chacune d'elles dans le modèle étudié.

L'adoption de ce type de modélisation est plus délicate que celle d'une matrice de transition. La dépendance aux conditions initiales est grande. Il est donc nécessaire de surveiller l'enchaînement des équations et les effectifs concernés à chaque étape, sous peine de dérive dans les reports présents (monopolisation d'une catégorie, disparition d'autres catégories...).

Prenons par exemple l'équation logistique définie pour déterminer le nombre de personnes inactives à chaque trimestre de la projection ; si la modélisation de cette équation n'est pas adaptée, le nombre d'inactifs générés au départ de la projection sera soit trop fort, soit trop faible, et cet effet ira crescendo au cours de la projection pour générer un nombre d'inactifs au final complètement irréaliste. Et même si les modèles logistiques qui suivent, simulant l'activité au Régime Général, dans un autre régime, le chômage, la maladie et l'invalidité, sont appropriés, le nombre de personnes actives passant par ces équations n'étant pas correct, les effectifs pour les transitions suivantes en seront affectés.

Ainsi, si l'un des modèles dichotomiques n'est pas ajusté, les enchaînements de transitions en aval de celui-ci risquent d'engendrer de mauvaises répartitions d'effectifs.

Outre ces difficultés, les temps de traitement pour faire tourner le modèle sont pour l'instant décuplés et les calages afin de respecter les hypothèses posées avant la projection sont plus difficiles à effectuer. C'est pourquoi le modèle Prisme repose pour l'instant sur la matrice de transitions décrite auparavant en attendant la fin des travaux sur les équations logistiques.

3.2 Modélisation des salaires

Le module salaire permet d'attribuer en projection un salaire annuel aux individus pour lesquels un report trimestriel d'activité au Régime Général est estimé.

Le calcul des salaires futurs est crucial : il détermine à la fois les masses de cotisations courantes et en partie les droits de retraite futurs.

Le salaire estimé est un salaire théorique annuel qui sera ensuite proratisé en fonction du nombre de reports de trimestres Régime Général enregistrés au compte pour obtenir le salaire déplafonné.

L'hypothèse d'évolution du Salaire Moyen Par Tête (SMPT) est appliquée sur les salaires déplafonnés à partir desquels les salaires plafonnés sont déduits (par comparaison au plafond de la Sécurité Sociale de l'année), sans qu'aucune distinction d'évolution des salaires selon le sexe ne soit posée.

Pour estimer les salaires, une régression MCO log-linéarisée est réalisée. Cette équation, inspirée de la modélisation de MINCER, se décline pour 4 publics (hommes ou femmes, avant ou après l'âge de fin d'études). L'estimation des équations est réalisée à partir des salaires non limités au plafond de la sécurité sociale (salaires déplafonnés), observés sur la période 1999-2004, issus de notre échantillon.

La formulation des équations est la suivante :

$$\ln(w_{jt}) = \alpha + X_{jt}\beta + u_j \quad \forall j, t \text{ (individu } j, \text{ période } t)$$

où α est la constante et $X_{j,t}$ le vecteur de variables explicatives

Les variables utilisées pour la modélisation des salaires avant l'âge de fin d'études sont l'âge courant, l'âge courant au carré ainsi qu'une indicatrice de présence de l'individu dans l'échantillon en $t-1$ ⁷.

Après l'âge de fin d'études, les variables explicatives de l'équation sont cet âge de fin d'études, l'expérience sur le marché du travail (issue de la différence entre l'âge courant et l'âge de fin d'études), l'expérience au carré, une indicatrice instantanée sur le chômage, une indicatrice permanente indiquant la présence d'une période de chômage dans la carrière, une indicatrice sur la présence d'une période d'invalidité, une indicatrice de présence dans l'échantillon en $t-1$, le pays de naissance, l'année de naissance de l'individu et le nombre d'enfants⁸.

L'estimation porte sur l'ensemble des salaires individuels, à temps partiel ou à temps complet, c'est-à-dire regroupant des situations très différentes. Le pouvoir explicatif du modèle est donc plus faible que celui des estimations d'équations de gains traditionnelles qui portent sur des salaires à temps complet uniquement. Toutefois, les coefficients sont relativement proches de ceux des autres études réalisées sur données françaises.

La structure par décile des écarts entre salaire observé et salaire estimé sur l'année 2004 est reproduite en projection afin d'assurer une diversité salariale au sein d'une même catégorie d'assurés (définie par sexe, par pays de naissance et par tranche d'âge). Les individus utilisés pour l'estimation conservent au cours de leur carrière projetée le résidu ainsi calculé.

Pour les individus connus avant projection pour lesquels nous ne disposons pas de salaire cette année-là, ainsi que pour les individus qui ont des reports de salaires au cours de leurs études, le résidu est normé.

4 La modélisation des départs en retraite

Le module « départ » a pour but de déterminer à chaque trimestre si un individu va partir en retraite, cela se traduisant dans le modèle par l'attribution de l'année et du trimestre de départ.

Les assurés rentrent dans le module de départ uniquement s'ils ont un âge compris entre 56 et 69 ans. Nous attribuons une date de départ fictive à 60 ans ou 65 ans aux individus qui n'ont pas acquis de droit au Régime Général, en fonction des autres régimes de passage connus.

Le but est de les exclure par la suite du modèle, pour ne plus en projeter la carrière et réduire le temps de computation.

Pour les assurés relevant du Régime Général, le module de départ est divisé en plusieurs parties qui s'appliquent en fonction de l'âge de l'assuré : la partie « non-réclamés⁹ », la partie

⁷ Indicatrice destinée à prendre en compte les effets de sélection liés à l'utilisation d'un panel non cylindré.

⁸ Uniquement pour les femmes.

⁹ Les non réclamés sont des individus ayant des droits acquis au Régime Général mais qui ne réclament pas leur pension de retraite. Il peut s'agir par exemple de personnes qui prolongent leur activité bien après 60 ans et qui meurent avant d'avoir cessé celle-ci, ou encore de personnes qui ont exercé une activité relevant du Régime Général, durant leur jeunesse pendant une durée limitée et qui ne pensent pas ou ne veulent pas réclamer leur pension (cas fréquents chez les nés à l'étranger)

« départs avant 60 ans¹⁰ », la partie « inaptitude¹¹ » et la partie « départs à partir de 60 ans ». En terme de programmation, chacune de ces parties fait l'objet d'une sous-macro qui est appelée dans le programme principal lorsque l'assuré est dans la tranche d'âge correspondante et n'a pas encore liquidé sa pension.

Les évènements simulés dans le module de départ reposent sur des équations logistiques. Pour estimer ces équations, une base a été constituée qui rassemble à la fois les individus déjà prestataires (ayant déjà liquidé leur pension) ainsi que les assurés qui n'ont pas encore pris leur retraite. Cette base d'estimation contient à la fois des informations constatées et des informations estimées (pour les assurés n'ayant pas encore pris leur retraite). Il aurait été possible d'estimer les équations sur la seule base d'informations constatées mais cela impliquait de travailler sur des générations anciennes. Le choix de travailler sur des générations plus récentes est lié au souhait de reproduire les tendances de comportements de départ les plus récentes, ce qui implique une base dans laquelle certaines informations sont partiellement estimées.

Nous allons présenter succinctement la modélisation des non-réclamés, des départs avant 60 ans et de l'inaptitude, puis nous présenterons plus en détail, à titre d'illustration de la mécanique utilisée dans le modèle, les départs à partir de 60 ans.

4.1 Les pensions des non-réclamés

Dans la chronologie de la projection, la détermination du caractère « non réclamé » d'un individu intervient le trimestre de ses 69 ans. Le modèle utilisé est un modèle logistique différencié selon le sexe (deux équations sont estimées séparément pour les hommes et pour les femmes). Il est estimé sur les assurés non-prestataires âgés de 69 ans ou plus des générations 1925 à 1927 (générations nombreuses et dont les assurés non-prestataires peuvent être considérés comme définitivement « non réclamés »).

Parmi les variables explicatives figurent la durée d'assurance validée au Régime Général, des indicatrices sur l'inactivité de l'assuré le trimestre précédant ses 69 ans et sur le caractère polycotisant de l'assuré, le pays de naissance, la moyenne des 30 meilleurs salaires annuels rapportée au montant du plafond de la Sécurité sociale...

Dans la mécanique du modèle, lorsque l'assuré atteint l'âge de 69 ans sans être prestataire, une sous-macro est appelée dans laquelle va s'effectuer la décision de le classer en non réclamé. Dans cette sous-macro, on calcule, pour chaque assuré, la probabilité qu'il soit non réclamé, à l'aide des coefficients estimés dans les deux modèles logistiques. Cette probabilité est ensuite comparée à un aléa : si l'aléa est inférieur, alors la pension de l'assuré devient non réclamée. Dans le cas contraire, l'assuré va automatiquement liquider sa pension puisqu'il est âgé de 69 ans.

4.2 Les départs avant 60 ans (ou départs en retraite-anticipée)

¹⁰ Depuis 2004, les assurés ont la possibilité de partir à la retraite avant 60 ans dans le cadre d'un départ anticipé. Sont éligibles à ce dispositif les assurés ayant eu une longue carrière ainsi que les assurés handicapés. Cette dernière catégorie représente environ 1000 départs par an, mais n'est pas simulée dans Prisme.

¹¹ La simulation de l'inaptitude n'est effectuée que pour les assurés âgés de moins de 65 ans. La simulation de l'inaptitude entraîne un départ immédiat à la retraite.

La retraite anticipée devait être modélisée de façon très précise, en raison de l'influence de cette mesure sur les enjeux de pilotage à court terme du régime.

Si l'assuré est éligible à un départ anticipé (s'il satisfait les conditions nécessaires à un départ avant 60 ans), le programme principal fait appel à la sous-macro des départs anticipés dans laquelle va être calculée une probabilité de liquider la pension.

Cette probabilité est calculée à l'aide de régressions logistiques estimées sur les prestataires ayant bénéficié d'une retraite anticipée, ainsi que sur les assurés qui réunissaient les conditions nécessaires. La modélisation est effectuée par âge exact (au trimestre près), et sexe, soit 32 équations.

Parmi les variables explicatives figurent la durée en trimestres depuis laquelle l'assuré remplit les conditions pour un départ anticipé, la situation de l'assuré vis-à-vis de l'emploi lors du trimestre précédent, la durée d'assurance validée au Régime Général, une variable binaire représentant le caractère polycotisant de l'assuré, le pays de naissance, une indicatrice permettant de savoir si l'assuré a connu une période de chômage (lui ayant permis de valider au moins un trimestre) au cours des 12 derniers mois, la moyenne des 30 meilleurs salaires annuels rapportée au montant du plafond de la Sécurité sociale...

En projection, la probabilité d'un départ en retraite anticipée est comparée à un aléa : si l'aléa est inférieur à la probabilité calculée alors il y a survenue de l'évènement « départ à la retraite ». Dans le cas contraire, rien ne se passe et l'on sort de la sous-macro.

Lorsque la décision de départ à la retraite est prise, une date de départ est attribuée, puis 3 autres sous-macro sont appelées. La première permet la mise à jour des durées cumulées (et donc l'attribution finale des durées de carrière), la deuxième met à jour les variables de salaire, la troisième donne lieu au calcul de la pension. Une fois ces traitements effectués, on revient dans le programme principal.

4.3 Les départs à partir de 60 ans

Lorsqu'un assuré n'est pas encore prestataire et qu'il est âgé d'au moins 60 ans, on va décider à chaque trimestre si son départ en retraite a lieu à cette date. Cette partie du module de départ est composée de deux étapes : dans un premier temps, on simule l'inaptitude au travail entraînant un départ immédiat, puis dans un deuxième temps et pour les autres, on simule la décision de départ en retraite.

a) L'inaptitude

La simulation de l'inaptitude n'est effectuée que pour les assurés de moins de 65 ans. La simulation du caractère inapte d'un assuré va déclencher un départ en retraite immédiat. L'estimation de l'inaptitude repose sur un modèle logistique, différencié en 6 équations selon le sexe et l'âge (3 classes d'âge : 60 ans, 60,25 ans et 60,5 ans et plus). Parmi les variables explicatives figurent la différence entre la durée requise pour l'obtention d'une pension à taux plein et la durée d'assurance totale de l'assuré, la moyenne des 30 meilleurs salaires annuels rapportée au montant du plafond de la Sécurité sociale, la situation vis-à-vis de l'emploi, le pays de naissance, l'âge de fin d'études, une indicatrice permettant de savoir si l'assuré a connu une période de chômage (lui ayant permis de valider au moins un trimestre) au cours des 12 derniers mois, une indicatrice permettant de savoir si l'assuré a connu une période de maladie (lui ayant permis de valider au moins un trimestre) au cours des 12 derniers mois, une variable binaire représentant le caractère polycotisant d'un assuré...

Ces variables finalement significatives correspondent au public supposé : faibles salaires, situation relativement précaire sur le marché du travail.

La simulation de l'inaptitude est liée à la comparaison entre un aléa et la probabilité calculée : si l'aléa est inférieur, l'assuré est déclaré inapte.

Une fois cette sous-macro exécutée, le module de départ fait appel à la sous-macro qui simule les départs en retraite à partir de 60 ans.

b) Les départs à partir de 60 ans

Dans cette sous-macro, on isole tout d'abord les assurés appartenant aux catégories d'inaptes ou d'invalides, ainsi que les assurés qui ont atteint 69 ans. Pour ces individus, la décision de départ en retraite est automatique, ils vont donc liquider leur pension.

Pour les individus restants, une probabilité de départ va être calculée. Celle-ci est basée sur plusieurs équations logistiques, dont l'estimation va être détaillée ici à titre d'illustration de la mécanique utilisée dans la modèle.

46 équations sont estimées (une pour chaque âge en trimestres compris entre 60 et 65,25 ans et une pour les âges supérieurs, pour les hommes et pour les femmes). Chacune de ces équations est estimée séparément en modifiant la liste de variables explicatives de manière à trouver la meilleure spécification possible (mesurée par les critères d'Akaike et de Schwartz et le taux de paires concordantes). Nous allons présenter ici deux exemples d'équations qui ont été estimées grâce à la PROC LOGISTIC sous SAS

La première concerne les départs en retraite des hommes de 60 ans :

```
proc logistic data=h60
    (where=(agexx=60)) descending outest=coef_deph1;
by agexx;
class dist_tx_plein_tr (ref='3') emp2 (ref='0') t (ref='4') pays_nais
(ref='1') / param=ref;
model liquid= dist_tx_plein_tr emp2 valid_rg poly chol mall tx_cot fin_etud
pays_nais sal_moy / selection=stepwise ;
title' deph1 ' ;
run;
```

La variable *liquid* (variable d'intérêt) est une variable binaire qui prend la valeur 1 lorsque l'assuré part en retraite lors du trimestre courant et 0 sinon. L'option 'descending' est spécifiée pour que SAS modélise la probabilité que *liquid* vale 1. Dans la clause 'class' sont spécifiées les variables catégorielles, avec la modalité de référence entre parenthèses.

Les variables explicatives sont les suivantes :

- *dist_tx_plein_tr* mesure en trimestres la différence entre la durée requise pour l'obtention d'une pension à taux plein et la durée d'assurance totale de l'assuré. Cette variable est découpée en 5 tranches : ($[<0]$; 0 ; [1-4] ; [5-19] ; $[>=20]$),
- *emp2* décrit la situation vis-à-vis de l'emploi. Elle compte trois modalités : '0' pour sans emploi, '1' pour emploi salarié et '2' pour emploi dans un autre régime,
- *valid_rg* est la durée d'assurance validée au Régime Général, c'est une variable numérique continue,
- *poly* est une variable binaire représentant le caractère polycotisant d'un assuré. Elle vaut 1 si l'assuré a validé des trimestres dans un autre régime,
- *chol* est une indicatrice permettant de savoir si l'assuré a connu une période de chômage (lui ayant permis de valider au moins un trimestre) au cours des 12 derniers mois,
- *mall* est une indicatrice permettant de savoir si l'assuré a connu une période de maladie (lui ayant permis de valider au moins un trimestre) au cours des 12 derniers mois,

- *tx_cot* est le ratio entre la durée de cotisation totale et la durée d'assurance totale,
- *fin_etud* est l'âge de fin d'études estimé dans le modèle Prisme,
- *pays_nais* est une variable dichotomique prenant la valeur '1' lorsque l'assuré est né en France, et prenant la valeur '2' sinon,
- *sal_moy* est la moyenne des 30 meilleurs salaires annuels (ou moins si l'assuré n'en totalise pas 30) rapportée au montant annuel du plafond de la Sécurité sociale.

La modélisation du départ est fortement axée autour de la notion de distance au taux plein. Cette variable a été préférée à la simple durée totale car la durée requise pour l'obtention d'une pension au taux plein est amenée à évoluer dans les années à venir. De plus, elle a été découpée en tranches car on ne peut conclure à une relation linéaire entre celle-ci et le fait de partir en retraite, notamment lorsque l'on considère les individus ayant une distance négative (qui peuvent avoir une incitation à prolonger leur activité pour bénéficier de la surcote).

Le modèle qui suit correspond à celui estimé pour les femmes dont l'âge est compris entre 65,5 ans et 68,75 ans et toujours non-prestataires à 65 ans. Il s'agit là de la dernière classe d'âge modélisée. Elle représentait environ 20 000 départs (hommes et femmes confondus) en 2005, soit 3,5% du flux de départs total. Malgré la grande taille de notre échantillon (1/20^{ème} de la population totale), nous n'avons donc pu estimer une équation pour chacun des âges exacts. En outre, il apparaissait inopportun de consacrer du temps à une procédure fastidieuse (28 équations à estimer) en comparaison avec le poids de cette population dans les liquidations totales.

```
proc logistic data=f65
    (where=(agexx>=65.5)) descending outest=coef_dep23;
class dist_tx_plein_tr (ref='0') t (ref='1') emp2 (ref='0') pays_nais
(ref='1')/ param=ref;
model liquid=tx_p v_rg agexx tx_cot t emp2 poly fin_etud /
selection=stepwise;
title' dep23 ' ;
run;
```

Dans cette équation logistique, de nouvelles variables explicatives apparaissent :

- *tx_p* est une indicatrice valant 1 si l'assuré a une durée d'assurance totale au moins égale à celle requise pour l'obtention d'une pension au taux plein¹²,
- *agexx* est l'âge exact (au trimestre près) de l'assuré, qui est ici introduit en tant que variable explicative puisqu'il peut prendre plusieurs valeurs. C'est une variable numérique continue.

Cette population de personnes âgées de plus de 65 ans et qui n'ont pas encore liquidé est relativement hétérogène. En moyenne, ces assurés sont très peu souvent en activité, et ils ont des durées relativement faibles. Dans cette population, on compte également une catégorie de personnes qui ne vont jamais liquider leur pension (qui seront classés dans la catégorie des « non-réclamés » dans notre modélisation). Ces individus tirent vers le bas les durées moyennes ainsi que les taux d'activité moyens. Parmi tous les assurés qui potentiellement peuvent liquider leur retraite à partir 65,5 ans, ce sont les variables d'activité, de durée validée au Régime Général ou encore le critère polypensionné qui vont être déterminants dans le choix des assurés qui liquident leur pension.

¹² A noter que dans cet exemple, toutes les personnes bénéficient d'une pension à taux plein puisqu'elles sont âgées de plus de 65 ans. Cette variable n'est en fait qu'une variable de durée comportant 2 classes.

Performances prédictives des équations logistiques sur les départs après 60 ans :

	Taux de bien classés	Sensibilité	Spécificité
Hommes	92,2%	37,5%	95,8%
Femmes	93,2%	40,0%	96,4%
Ensemble	92,7%	38,7%	96,1%

Dans le tableau ci-dessus, nous pouvons constater que le taux de bien-classés est de l'ordre de 93% avec une sensibilité valant près de 39% et une spécificité de 96%.

Dans la programmation de Prisme, la sous-macro assurant la gestion des départs en retraite après 60 ans calcule donc une probabilité de départ pour les individus non inaptes et non invalides, et dont l'âge est compris entre 60 ans et 68,75 ans. Cette probabilité est calculée à l'aide des coefficients estimés dans les 46 équations logistiques. Au niveau individuel, la probabilité est ensuite comparée à un aléa : lorsque ce dernier est inférieur, il y a survenue du départ en retraite. Dans ce cas, il y a donc attribution d'une date de départ puis, on attribue une valeur finale aux durées de carrière, on complète les variables de salaire et enfin, on procède au calcul de la pension.

5 La modélisation de la réversion

La législation française en matière de retraite prévoit l'attribution d'une pension de réversion aux veufs ou veuves d'assurés du régime général dont les ressources ne dépassent pas un certain montant. Le montant de la pension éventuellement attribué constitue ainsi un peu plus de la moitié du montant de retraite que la personne décédée percevait ou aurait perçu par le régime de retraite.

Afin de calculer les droits dérivés, le modèle de projection intègre un module réversion, lequel n'intervient qu'une fois les droits propres calculés.

Lorsque la personne est déjà prestataire d'un droit dérivé au point de départ de la projection, elle le reste jusqu'à son décès, sauf dans le cas de nouveaux événements comme un nouvel emploi ou l'attribution d'une nouvelle pension.

Par contre, dans le cas de nouveaux prestataires d'une pension de réversion en cours de projection, il faut en amont définir la population de bénéficiaires potentiels. Pour cette raison, le module de réversion se compose de plusieurs étapes, que nous allons détailler.

La population du modèle est constituée d'individus sans liens familiaux.

Notre première étape consiste donc à créer des ménages parmi les individus présents dans la table. Il n'y a donc pas d'assurés mariés avec des conjoints extérieurs à la table, laquelle rassemble les individus nés en France et/ou ayant acquis un droit au titre de la protection sociale en France¹³.

Nous posons une hypothèse de taux de mariage de 90% pour les générations antérieures à 1950, puis une baisse progressive de ce taux pour atteindre 75% à partir de la génération 1965, sur la base du taux de personnes mariées à 50 ans calculé par l'INSEE.

¹³ En conséquence, en projection, nous ne marions pas d'individus connus dans cette base à des personnes résidentes à l'étranger et inconnues du système de protection sociale français. Les veuves de l'étranger sont donc "remplacées" par des veuves résidant en France dans notre modèle.

Le mariage se réalise par rapprochement de deux assurés de sexe différent selon une différence d'âge et selon une différence de niveau d'études. Ces différences ont été fixées sur la base des données de l'enquête Famille et d'analyses réalisées en collaboration avec l'INED. Le choix du conjoint se fait par tirage aléatoire parmi les candidats possibles au mariage, correspondant à une classe relative de niveau d'études et à un âge visés par l'assuré.

Concrètement, les conjoints et conjointes potentiels sont réunis et triés, les conjointes choisissent ou non de s'unir (par tirage aléatoire) avec les hommes qui leur sont proposés, jusqu'à ce que le taux de mariage de cette catégorie d'union soit atteint. Les célibataires restant sont remisés et éventuellement sélectionnés à nouveau pour une nouvelle configuration d'écart d'âge et d'âge de fin études.

Différentes combinaisons de différences d'âge et de niveau d'études ont été testées afin d'obtenir une proportion d'individus mariés par génération conforme aux données fournies par l'INED, et de reproduire les différences d'écart d'âge constatées entre conjoints.

En fonction des décès projetés de l'un des deux conjoints, une population en situation de veuvage se forme au fur et à mesure de la projection.

Dans un second temps, nous déterminons l'éventuel droit à la réversion des conjoints survivants en observant les caractéristiques de leurs carrières ainsi que celle de leurs conjoints décédés.

Le calcul d'un nouveau droit de réversion nécessite d'estimer les ressources des bénéficiaires potentiels, puisque seuls les veufs ou veuves dont les ressources sont inférieures à un certain seuil peuvent bénéficier d'une pension de réversion du régime général. Dans l'estimation des ressources sont cumulés les salaires, les pensions de droit propre tous régimes, les revenus du patrimoine, et éventuellement les revenus du nouveau conjoint en cas de remariage de l'assuré survivant.

Mise à part la pension de droit propre versée par le régime général, les autres éléments de revenus ne sont pas connus dans les données de la Cnav et font donc l'objet d'une estimation, laquelle dépend de la situation professionnelle du conjoint survivant.

Pour les actifs, nous estimons un revenu d'activité (ou un revenu de remplacement) à partir des salaires calculés pour les salariés du secteur privé, selon le genre et l'âge.

Pour les personnes en retraite, nous estimons leur pension tous régimes, ainsi que les ressources du patrimoine détenues en propre. La pension tous régimes des prestataires qui ont une retraite de droit propre au régime général est estimée à partir du montant de retraite de base versé par ce dernier, et des durées d'assurance respectives effectuées dans les différents régimes, et en fonction des résultats extraits de l'échantillon inter-régimes (EIR)¹⁴.

A partir de travaux effectués par l'INSEE ou d'autres organismes de recherche¹⁵, nous estimons les revenus du patrimoine en fonction du décile auquel le montant de la pension appartient.

Pour les prestataires bénéficiant d'une pension totale d'un montant compris entre le 7^{ème} et le 9^{ème} décile de la distribution des pensions, le revenu du patrimoine représente 25% de la pension totale. Pour les assurés qui ont une pension supérieure au 9^{ème} décile, le revenu du patrimoine est estimé à 50% de la pension ; pour les prestataires dont le montant de la retraite est inférieur au 7^{ème} décile, le revenu du patrimoine est considéré comme nul.

¹⁴ L'EIR est un échantillon constitué tous les quatre ans à partir des données fournies par la quasi-totalité des organismes de retraite obligatoire ; il permet de reconstituer la retraite totale des retraités.

¹⁵ Enquête emploi, enquête patrimoine 1998, travaux de la DREES...

Une fois la population des bénéficiaires de la réversion constituée (soit la population en situation de veuvage vivant l'année N, et dont les ressources sont inférieures à un certain seuil), nous calculons, pour chaque année de projection, le montant des droits dérivés.

C- Déroulement type et résultats

1 Déroulement type

Le modèle Prisme, qui requiert des ressources informatiques importantes, est exécuté sur un serveur Unix, serveur quadri-processeurs disposant de 8 Go de mémoire vive. Ce serveur est relié à des baies de stockage d'une capacité totale de 2,4 To.

Il est entièrement développé en macro langage Sas. Il est composé de 2 macros principales (une pour les droits propres et une pour les droits dérivés) et de 22 sous-macros. Dans sa version actuelle, Prisme représente environ 6 200 lignes de programme.

Les macros sont réparties en 4 thèmes :

- démographie : 5 (naissances, immigration,...) ;
- carrière : 6 (transitions, salaires, durées validées,...) ;
- retraite : 7 (inaptitude, départs, calcul de la pension,...) ;
- réversion : 4 (ressources, calcul de la pension de réversion,...).

Prisme nécessite beaucoup de ressources informatiques. Une optimisation de la programmation et du déroulement de la projection est indispensable pour obtenir des temps de traitement acceptables. Ces optimisations portent sur 4 axes :

- la limitation des données en lecture et écriture : dans l'estimation des événements, Prisme a besoin d'un historique de carrière d'une année (4 trimestres). Pour éviter la lecture et l'écriture d'un volume d'informations inutiles à la projection, à la fin de chaque année de projection, les données de carrière relatives à l'année précédente sont stockées dans une table distincte. En fin de projection, l'ensemble des tables de stockage sont appariées pour obtenir une table contenant l'intégralité des carrières projetées.
- la génération de nombres aléatoires : chaque événement (naissance, décès, transition de carrière, retraite,..) est déterminé par la comparaison d'une probabilité (logit ou matrice de probabilités) avec un nombre aléatoire.
Pour une projection au 1/20ème à l'horizon 2050, ce sont environ 8 milliards d'événements qui seront estimés, et donc autant de nombres aléatoires qui seront nécessaires.
Les nombres aléatoires générés doivent répondre à une contrainte : garder d'une variante de projection à l'autre exactement les mêmes aléas pour chaque individu pour écarter les effets purement aléatoires dans l'analyse des variantes.
Compte tenu du nombre d'aléas nécessaires, il est impossible d'enregistrer toutes les séries d'aléas pour l'ensemble des individus.
La solution est de générer les aléas de manière contrôlée au cours de la projection avec la routine ranuni de Sas (aléa issu d'une distribution uniforme). Le contrôle de la production des aléas se fait grâce à l'utilisation de germes : à partir d'un germe unique, ranuni générera la même série de nombres aléatoires.
Dans Prisme, sont créées des classes d'individus à l'intérieur desquelles l'ordre des individus est identique quelle que soit la variante (notamment celle démographique, qui peut changer le nombre d'individus).

- la création de formats : au cours de la projection, pour chaque événement, la modélisation requiert dans l'étape data principale des données de type barème (plafond de la sécurité sociale, montant minimum de pension, durée taux plein, etc.) et des données déterminées avant projection (matrice de probabilités, coefficients de logit, etc.).
Ces données sont stockées dans des tables Sas indépendantes. Pour faire appel à ces données nous utilisons des formats qui ont comme avantage un gain de temps d'exécution et également une facilité de mise en œuvre.
Prisme utilise plus de 240 formats. L'ensemble des barèmes, des coefficients de régression logistiques, etc. sont enregistrés en formats et sont appelés à l'aide de la fonction input.
- l'utilisation de MP-connect : une tâche Sas (par exemple une étape DATA) s'exécute sur 1 seul processeur. Donc, le modèle écrit 'classiquement' n'utiliserait que 25% de la puissance du serveur Unix (quadri-processeurs). Pour y remédier, le modèle a été écrit de manière à être exécuté simultanément sur 2 processeurs : la population est répartie à chaque pas annuel entre les moins de 50 ans et les plus de 50 ans. C'est le principe du MP-connect (MP pour Multi-processors) : à l'intérieur de la session Unix principale (lancement du modèle), N sous sessions Unix sont créées. Les traitements sont répartis entre les sous-sessions et sont exécutés en parallèle sur N processeurs.

Lorsqu'un exercice de projection est lancé, plusieurs paramètres sont renseignés. Ces paramètres, pour la partie droits directs sont intégrés dans l'appel de la macro principale.

Les paramètres modifiables sont, entre autres, les suivants :

- le numéro de projection qui permet de changer la série des aléas générés et donc de mesurer la sensibilité de la projection liée aux aléas (de 1 à 5)
- l'indice conjoncturel de fécondité (1,5 ; 1,8 ; 1,9 ; 2,1),
- l'hypothèse d'immigration (faible, moyenne, haute),
- le taux d'échantillonnage retenu ($1/20^{\text{ème}}$, $1/100^{\text{ème}}$, $1/1000^{\text{ème}}$, $1/10000^{\text{ème}}$),
- l'année de début de projection,
- l'année de fin de projection,
- la législation retenue (Balladur = législation après réforme de 1993 mais sans la réforme de 2003; Fillon = législation après réforme de 2003; Sans = législation avant réformes de 1993),
- le type de quotient de mortalité utilisé (insee = quotients de mortalité INSEE appliqués à tout le monde ; cnav = quotients de mortalité INSEE appliqué aux non retraités et quotients de mortalité CNAV appliqués aux retraités).

En annexe est présenté un schéma détaillé illustrant la mécanique du modèle en programmation SAS.

Les paramètres de projection du modèle pour la partie droits dérivés sont intégrés dans l'appel de la macro principale. Ces paramètres sont, entre autres, les suivants :

- le plafond de ressources pour une personne seule,
- le minimum de pension de réversion.

Actuellement le temps de traitement pour faire tourner le modèle au $1/20^{\text{ème}}$ est de l'ordre de cinq heures. Au $1/100^{\text{ème}}$ il est d'une heure.

A l'issue de la projection, des sorties automatiques sont exécutées. Un certain nombre d'indicateurs agrégés sont calculés en ce qui concerne le stock et les flux des retraités, comme les pensions moyennes, les masses de prestations et le nombre de départs à la retraite. Ces indicateurs peuvent éventuellement être déclinés par sexe, âge, type de pension, année de départ,... Des sorties concernant la carrière des assurés en projection sont également prévues.

En ce qui concerne les droits dérivés, le fichier généré regroupe les informations du flux et du stock. Il contient des données par année civile et par sexe concernant le flux (entrant, sortant, bénéficiaires âgés de moins de 55 ans,...) et le stock (stock de l'année, des vivants en fin d'année, des bénéficiaires âgés de moins de 55 ans, déclinaison des avantages complémentaires,...).

2 Résultats et simulations

2.1 Résultats – scénario central

Les résultats présentés ci-après correspondent à l'exercice de projection réalisé en 2007 à la demande du Conseil d'Orientation des Retraites (COR)¹⁶. Différentes séries d'hypothèses ont été successivement retenues pour ces projections. Les graphiques suivants sont relatifs au scénario central.

a) Le solde technique de la CNAV

Le solde technique de la CNAV présenté dans cette sous-partie correspond à la différence entre les cotisations, auxquelles s'ajoutent les transferts de l'Etat au titre des avantages de retraite relevant de la solidarité nationale (chômage et majorations pour enfants), et les prestations.

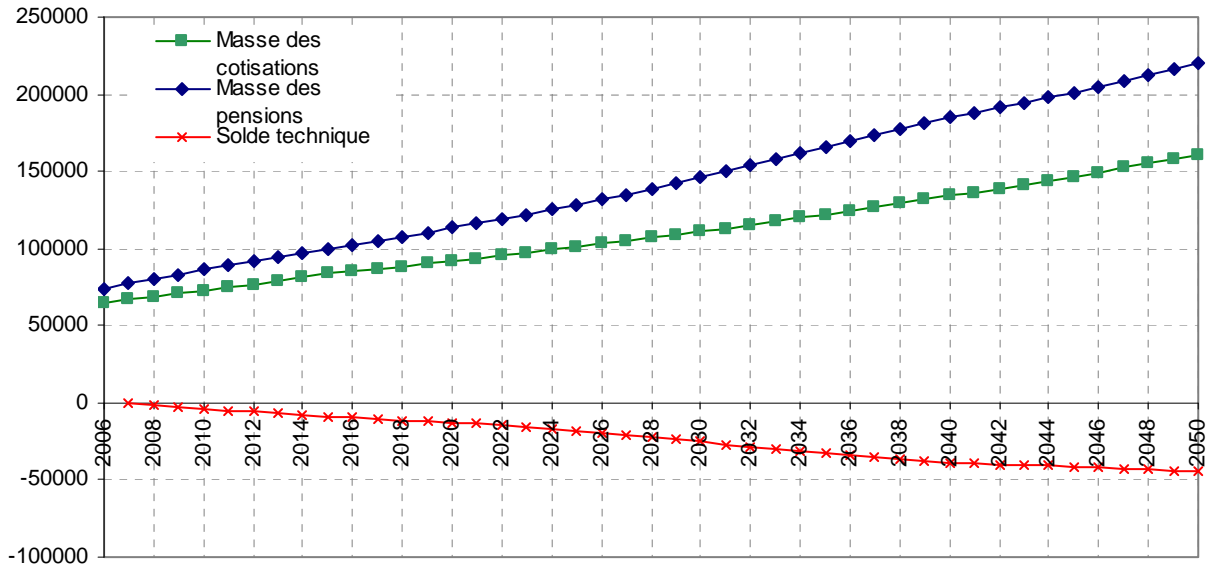
Côté ressources, la masse des cotisations devrait s'élever à 160 Md€ à l'horizon 2050, liée à la baisse progressive supposée du taux de chômage jusqu'en 2015 (où il atteindrait 4,5%) et à l'augmentation du nombre de cotisants sur la même période.

Côté dépenses, l'augmentation du nombre de prestataires conjuguée à l'évolution de la pension moyenne explique la hausse importante de la masse des pensions : la masse va plus que tripler sur la période 2006-2050.

Ces évolutions des ressources et dépenses ont pour conséquence un solde technique déficitaire de 45 Md€ à l'horizon 2050.

¹⁶ Le Conseil d'orientation des retraites est un lieu permanent d'études et de concertation entre les principaux acteurs du champ des retraites. Il suit l'évolution des régimes et fait des propositions pour assurer leur solidité financière et leur fonctionnement solidaire.

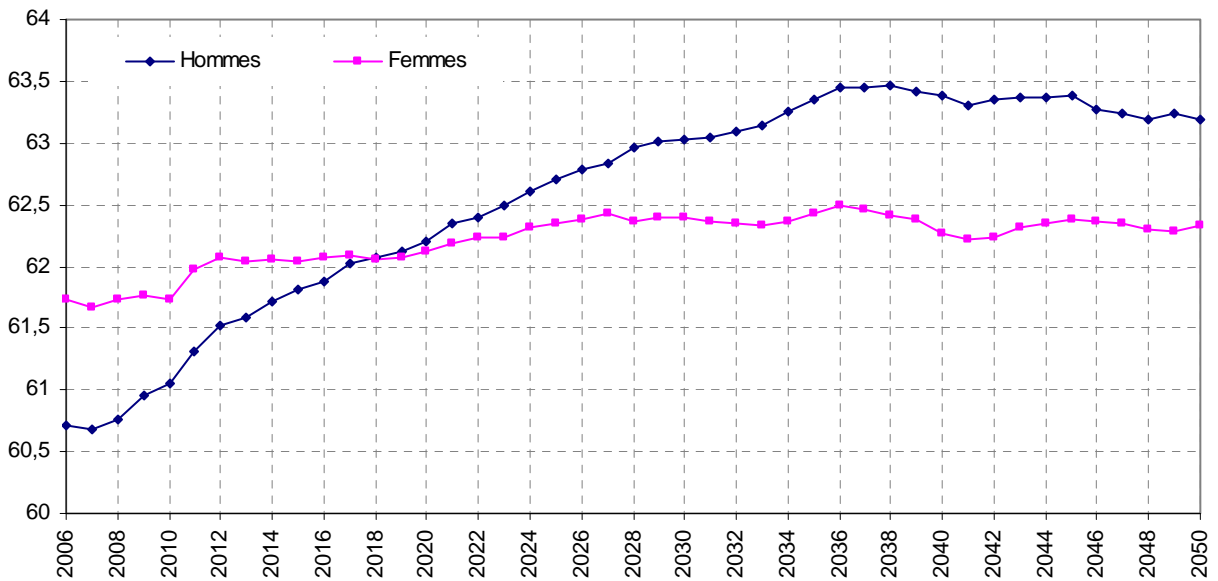
Solde technique (M€)



b) Age moyen du départ en retraite

Pour son scénario central, le COR a retenu l'hypothèse selon laquelle les mesures relatives à la surcote et à la décote inciteraient les hommes à reculer leur départ en moyenne de 0,6 an et les femmes à avancer de 0,3 an. Ces hypothèses, implémentées dans Prisme, expliquent l'évolution différente entre hommes et femmes notamment en terme d'âge moyen de départ à la retraite.

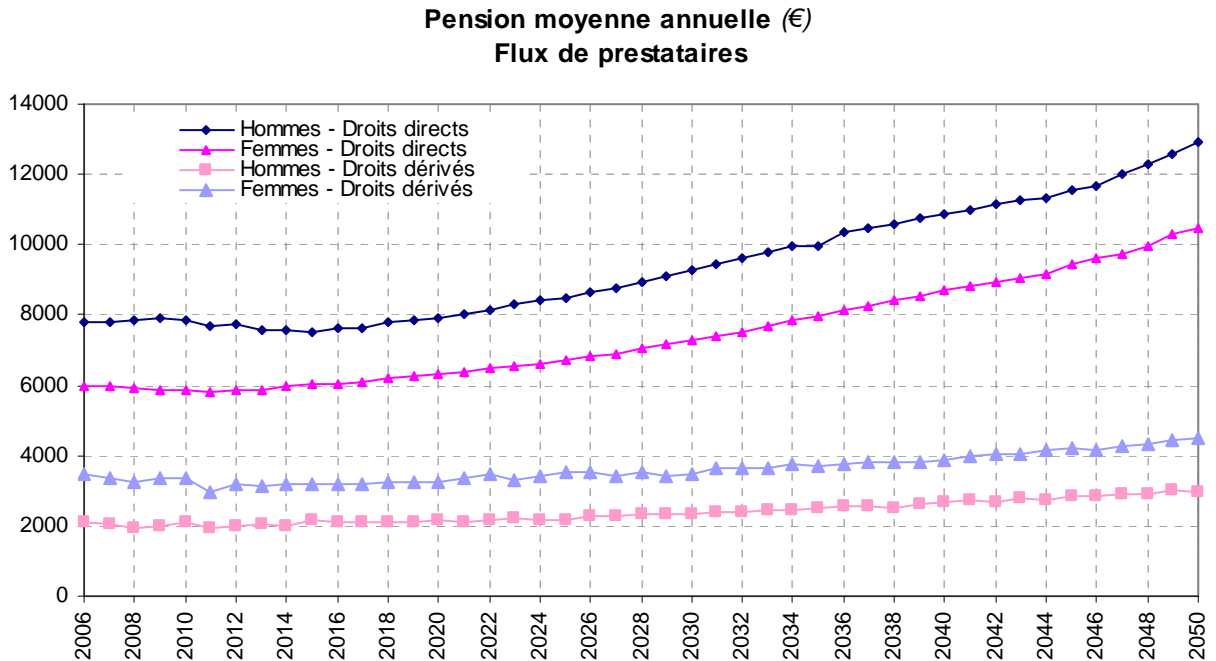
Age moyen du départ en retraite



c) Pension annuelle moyenne du flux de prestataires

L'évolution de la pension moyenne du flux de prestataires sur la période 2006-2050 est, respectivement, de 66% pour les hommes et de 75% pour les femmes (droits directs), et de 39% pour les hommes et 28% pour les femmes (droits dérivés).

Cette augmentation est liée à l'évolution du salaire moyen et des durées d'assurance sur la période de projection.

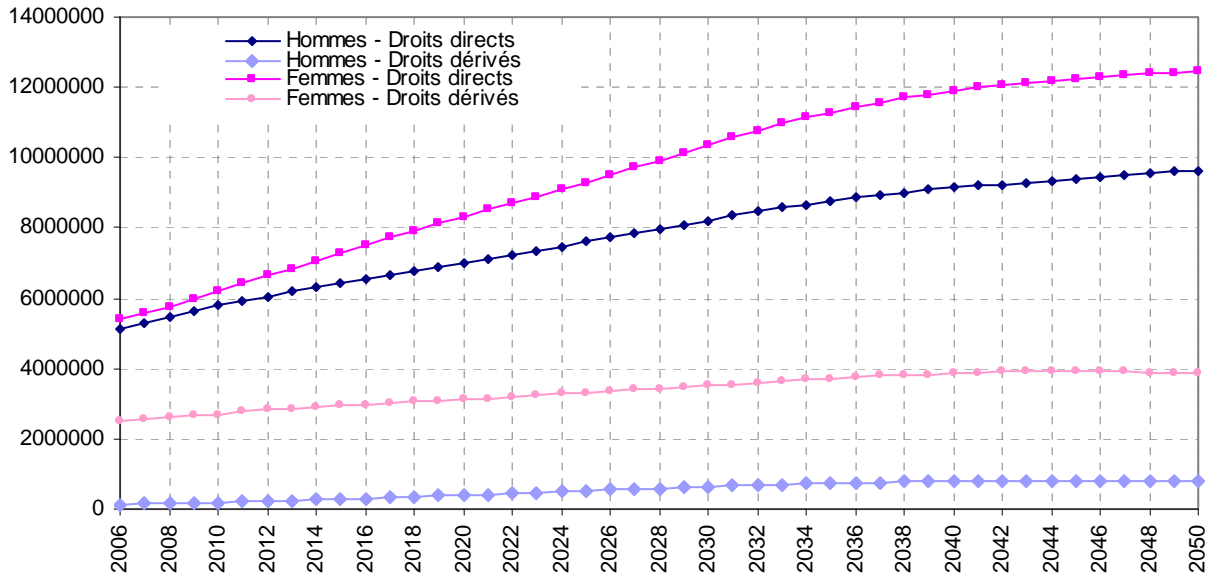


d) Stock de prestataires

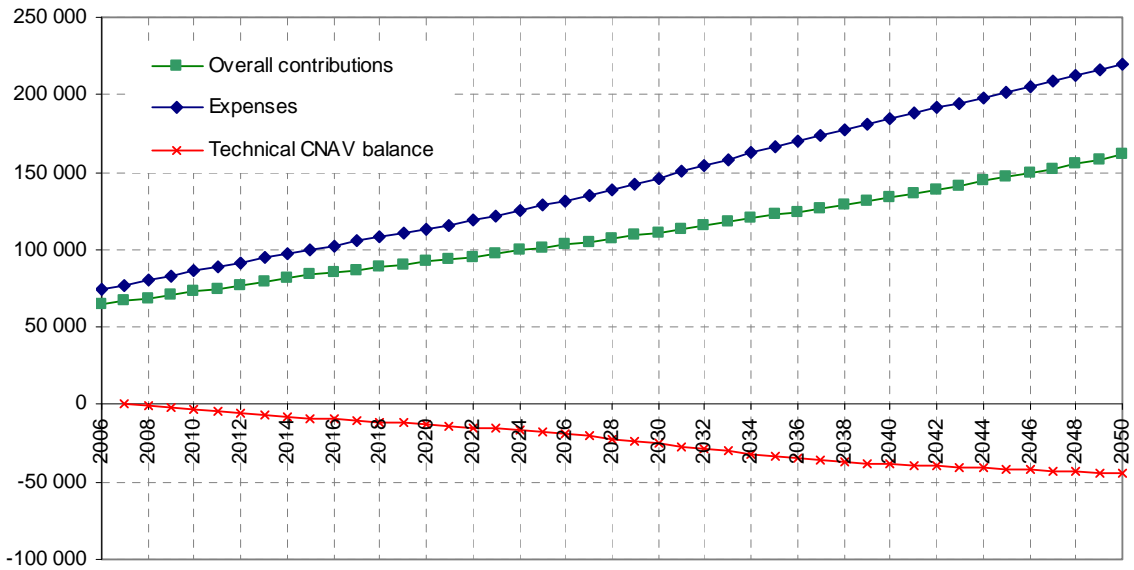
D'ici à 2050, le nombre de prestataires de droits directs sera multiplié par 2 pour les hommes et par 2,5 pour les femmes.

Le nombre de bénéficiaires de droits dérivés en stock augmentera régulièrement pour les hommes comme pour les femmes jusqu'en 2040 avant de se stabiliser.

Stock de prestataires



Technical CNAV balance (M€)



2.2 Simulation du décalage de l'âge légal de 60 à 61 ans¹⁷

La Cnav réalise également des simulations de changement de la législation, à la demande du COR, de la Direction de la sécurité sociale ou de certains ministères.

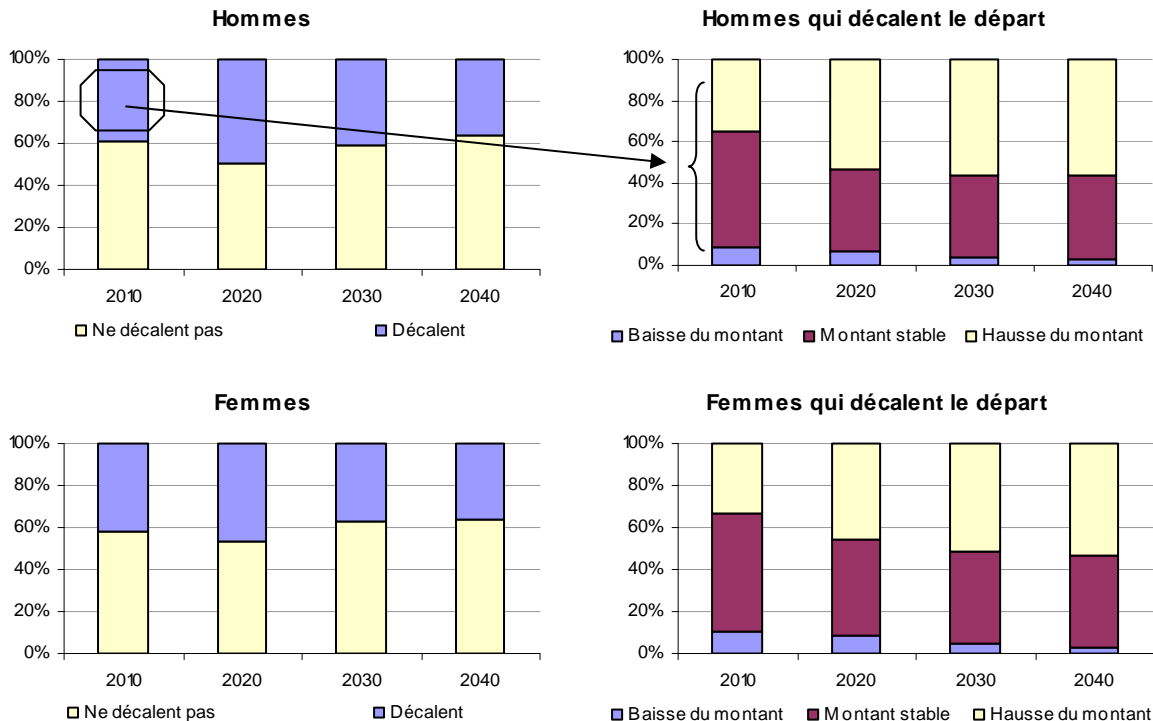
Une de ces demandes a par exemple porté sur le recul progressif de l'âge légal de la retraite de 60 à 61. Le relèvement serait appliqué de manière progressive : 1 trimestre par an entre 2009 et 2012, pour atteindre un âge légal cible de 61 ans.

Cette simulation nécessite la formulation d'un certain nombre d'hypothèses :

- Malgré le recul de l'âge légal, les bornes d'âge de retraite anticipée sont supposées invariantes : ainsi la mesure reste ouverte dès 56 ans.
- Entre 60 ans et le nouvel âge légal, un départ en retraite anticipée est possible aux conditions applicables à 59 ans.
- Le recul de l'âge légal est également appliqué aux assurés inaptes et invalides.

a) Pourcentage de prestataires concernés par la mesure

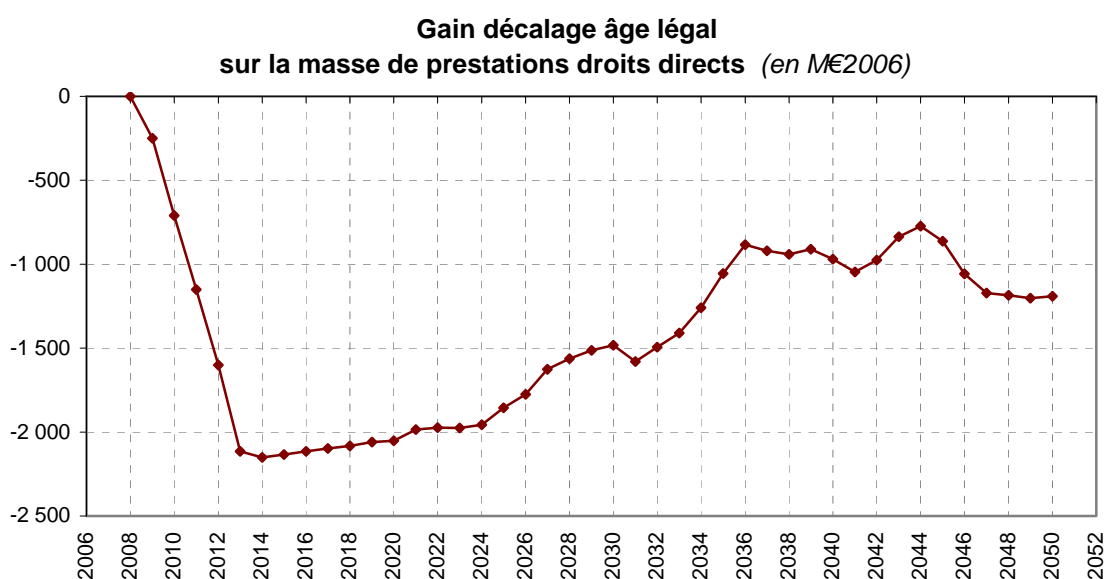
Chaque année, entre un tiers et la moitié des assurés décalerait son départ à la retraite. Parmi les prestataires qui décaleraient le départ, près de la moitié verrait le montant de la retraite augmenter, du fait de la poursuite de l'activité. L'augmentation des montants de pension concernerait notamment les hommes.



¹⁷ L'incidence sur les dépenses de prestations est mesurée ici uniquement pour les droits directs (l'impact d'un décalage d'âge légal sur les masses de droits dérivés devrait être très limité)

b) Incidence du recul de l'âge légal sur les masses de prestations de droits directs

Le décalage de l'âge légal de 60 à 61 ans réduirait les masses de prestations de près de 2 milliards d'euros à l'horizon 2020. Le gain se réduirait à partir de 2025 sous l'effet de l'augmentation des pensions moyennes et s'établirait à 1,2 Mds€ en 2050. En effet, les décalages permettraient aux assurés d'augmenter leurs droits en validant davantage de trimestres, et dans certains cas d'acquérir le taux plein ou encore en améliorant leur SAM pour ceux qui poursuivraient leur activité salariée et percevraient de nouveaux salaires.

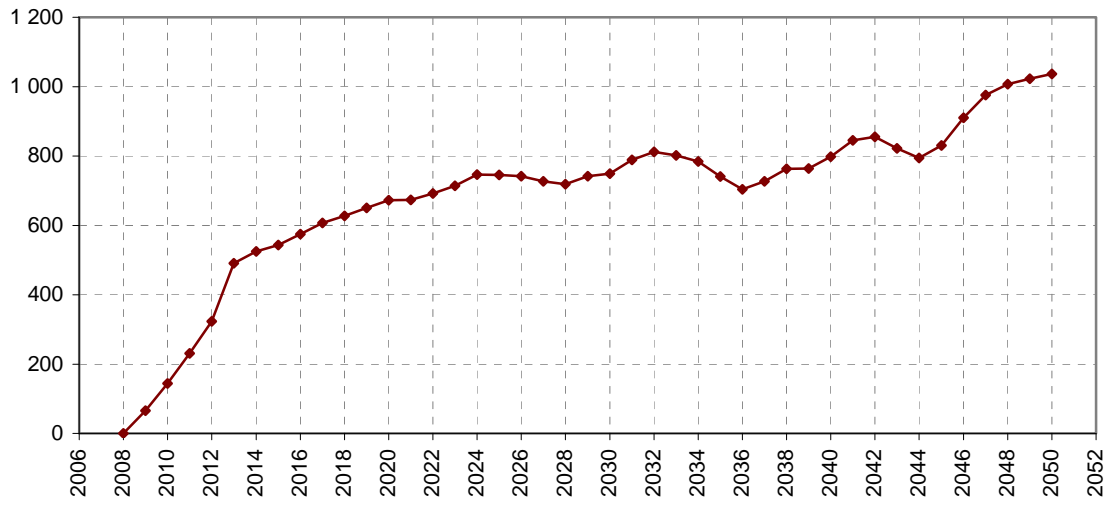


c) Incidence du recul de l'âge légal sur les masses de cotisations

Le recul de l'âge légal provoque des départs plus tardifs : lorsqu'un assuré décale son départ, il a été fait l'hypothèse que celui-ci reste, vis à vis de l'emploi, dans la même situation que celle précédant son report initial (dans la situation de référence avec l'âge légal à 60 ans).

En terme de masse de cotisations, uniquement pour le Régime Général, les cotisants supplémentaires engendreraient près de 700 millions d'euros de cotisations supplémentaires en 2020 et 1,04 Mds€ en 2050.

Suppléments de cotisations (en M€2006)



d) Incidence globale de la mesure

En ajoutant les effets de la mesure de décalage de l'âge légal à 61 ans sur les prestations aux effets sur les cotisations, le gain pour le Régime Général serait de 2,7 Mds€ en 2020 et de 2,2 Mds€ en 2050.

Conclusion – Avenir de *PRISME*

Le modèle *PRISME* répond actuellement aux évaluations de réformes potentielles demandées par les organismes concernés par la retraite en France (Ministères, Conseils...). Il est utilisé à la fois en tant qu'outil de projection de long terme et modèle de prévisions financières de très court terme, ce qui lui impose une gymnastique pas toujours évidente d'actualisations et de cadrages annuels qui ne remettent pas en question les solutions de long terme annoncées précédemment.

Son utilisation ne s'arrête pas là : il sert, par exemple, pour projeter les effets éventuels de modification radicale du fonctionnement même du régime, dans le cas d'un passage à un régime en point ou en comptes notionnels. Il se révèle utile pour évaluer les conséquences de réformes à l'intersection des domaines retraite, maladie et famille.

PRISME est devenu, à notre sens, une référence indispensable dans le processus de réforme continue des retraites tel qu'il semble se dessiner en France. D'autres articles viendront compléter à l'avenir cette présentation pour montrer l'étendue des possibles autorisée par ce modèle.

Annexe

Le schéma suivant illustre la mécanique du modèle. Il souligne l'utilisation, dans notre programmation, des *macros* (blocs d'instructions, externes au programme principal), des fonctions *symget* (fonctions qui permettent de récupérer des *macrovariables*), ainsi que des *formats* (données stockées dans des tables Sas indépendantes).

Ces fonctions permettent un gain de temps, comme souligné dans la partie C.

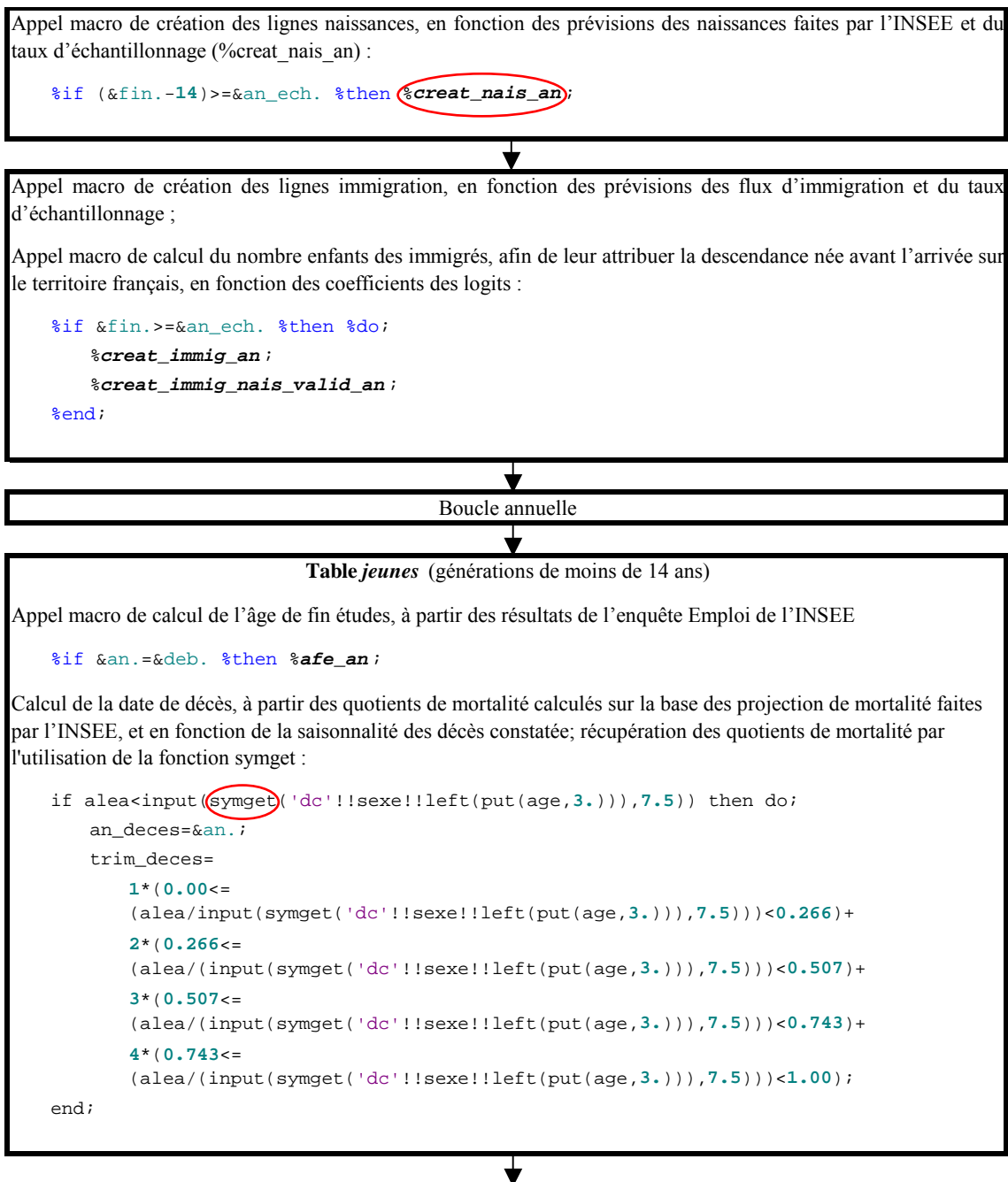


Table *actifs1* (générations en âge de travailler entre 14 et 49 ans)

Calcul décès (comme pour les *jeunes*)

Appel macro de calcul des naissances, afin d'incrémenter le cas échéant la descendance des femmes en âge fécond, en fonction des coefficients des logits :

```
if sexe = '2' and max(15,fin_etud)<=age<50 and nb_enf<5 and &an.>(&an_ech.-2) and
(an_deces=. or an_deces>&an.) then do;
  if nb_enf=. then nb_enf=0;
  %naissances_an ;
end;
```

Boucle trimestrielle

Affectation des types de reports validés et du nombre de trimestres correspondant

Appel macro de calcul des salaires, afin d'attribuer un salaire annuel aux individus avec un report d'activité salariée, en fonction des salaires observés sur le passé

```
...
if trim=4 then %salaires_an ;
```

Appel macro de mise à jour des durées cumulées ;

Appel macro de mise à jour des meilleurs salaires, pour le calcul du salaire annuel moyen (SAM) :

```
if (an_deces=. or an_deces>=&an.) and an_entree<=&an.
and an_ej=. and prem_rep^=. and age>=prem_rep then do;
  %maj_cumuls_an (tr=4) ;
  %maj_sal_max_an ;
end;
```

Table *actifs2* (générations en âge de travailler entre 50 et 70 ans)

Calcul décès, à partir des quotients de mortalité estimés sur la base des décès constatés dans les bases CNAV, par type de pension (normale, inapte, ex-invalide)

Boucle trimestrielle

Affectation des types de reports validés et du nombre de trimestres correspondant

Appel macro de calcul des salaires (comme pour les *actifs1*)

Appel macro de mise à jour des durées cumulées (comme pour les *actifs1*)

Appel macro de calcul des assurés non réclamés, afin de sortir de la boucle calculant la date de départ à la retraite les individus qui ne réclament par leur retraite

```
if %if &retro_dep.=0 %then &an. >= &an_ech. and ;
categ='' and agex=65 and trim_nais=trim then %non_reclames ;
```

Appel macro de calcul du départ à la retraite, afin d'attribuer le cas échéant une date de départ à la retraite, à

Appel macro de calcul des montants de retraite, afin de déterminer les montants de la pension des assurés qui partent à la retraite ; récupération de la durée taux plein par appel d'un format (&liq.taux.) :

```
if an_ej=. and agex>=60 then do;
  if categ='' and 60<=agex<65 then %inaptitude ;
```

```

%if &departs.=tx_plein %then %do; /* Départs taux plein */
  if an_ej=. and valid_tot_tr>=input(put(generation,&liq.tx.),3.) or
  agex=65 or categ^='') then do;
    an_ej=&an.; trim_ej=trim;
    %maj_cumuls_an(tr=(trim-1));
    %salaires_an;
    %calcul_pension_an;
  end;
%end;
%else %do; /* Départs avec probabilités */
  if an_ej=. then do;
    %if "%substr(&departs.,1,3)"="mat" %then %departs_matrice;
    %else %if "%substr(&departs.,1,3)"="log" %then %departs_logit;
    %else %departs;
  end;
%end;
end;

```

Appel macro de mise à jour des durées cumulées (comme pour les *actifs1*)

Appel macro de mise à jour des meilleurs salaires (comme pour les *actifs1*)

Table anciens (générations retraitées d'âge compris entre 71 et 115 ans)

Calcul décès (comme pour les *actifs2*)

Table décédés

Glossaire

Coefficient de proratisation : ratio rapportant la durée d'assurance au régime général à la durée de proratisation. La durée de proratisation est la durée nécessaire pour bénéficier d'une pension entière. Elle est de 150 trimestres (37,5 ans), pour les assurés nés avant 1944, et elle augmente progressivement au fil des générations, pour atteindre 164 trimestres (41,5 ans) pour les assurés nés à partir de 1952.

Conseil d'orientation des retraites (COR) : Le Conseil d'orientation des retraites est un lieu permanent d'études et de concertation entre les principaux acteurs du champ des retraites. Il suit l'évolution des régimes et fait des propositions pour assurer leur solidité financière et leur fonctionnement solidaire.

Décote : Réduction définitive appliquée au montant de la pension d'un assuré qui choisit de partir en retraite avant d'avoir atteint la durée de cotisation nécessaire (ou l'âge requis) pour bénéficier d'une pension de retraite à taux plein. Le taux de décote est de 1,25% par trimestre manquant au taux plein, pour les assurés nés avant 1944, et il baisse progressivement, pour atteindre 0,6875% par trimestre manquant, pour les assurés nés à partir de 1952.

Echantillons inter-régimes de cotisants (EIC) : La Direction de la recherche, des études, de l'évaluation et des statistiques (DREES), qui relève du Ministère de la Santé et des Sports, interroge tous les quatre ans la quasi-totalité des organismes de retraite obligatoire, de base et complémentaires. Ceux-ci fournissent des données sur les droits à la retraite acquis par un échantillon anonyme d'individus (durées validées, motif de validation, salaires portés au compte, points acquis, etc.). Le rapprochement, individu par individu, de ces informations, permet une connaissance globale des droits à la retraite acquis par les actifs, sans reconstitution à partir de données déclaratives.

Echantillons inter-régimes de retraités (EIR) : La Direction de la recherche, des études, de l'évaluation et des statistiques (DREES), qui relève du Ministère de la Santé et des Sports, interroge tous les quatre ans la quasi-totalité des organismes de retraite obligatoire, de base et complémentaires. Ceux-ci fournissent des données sur les avantages de retraite versés à un échantillon anonyme d'individus. Le rapprochement, individu par individu, des montants en provenance des différents régimes permet de reconstituer la retraite totale de chaque retraité.

Institut national d'études démographiques (INED) : L'Ined a été créé en 1945 avec pour mission « d'étudier les problèmes démographiques sous tous leurs aspects. À cet effet, l'institut rassemble la documentation utile, ouvre des enquêtes, procède à des expériences et suit les expériences effectuées à l'étranger ».

Institut national de la statistique et des études économiques (INSEE) : L'Insee a été créé en 1946 avec pour mission de collecter, produire, analyser et diffuser des informations sur l'économie et la société française. Il assure la coordination du service statistique public français.

Monopensionné : Titulaire d'avantages vieillesse ou invalidité dans le seul régime général.

Pension d'invalidité ou pour les ex-invalides : Pension accordée au taux plein, quelle que soit la durée d'assurance validée, à partir de 60 ans, à un assuré qui présente un handicap ou une incapacité de travail.

Pension de réversion : Pension versée au conjoint survivant d'un assuré qui avait acquis des droits à une pension de vieillesse. Dans le régime général des salariés la pension de réversion est fonction des ressources du conjoint survivant.

Période assimilée : Période d'interruption de travail assimilée à une période d'assurance pour l'ouverture du droit et le calcul de la pension. (maladie, maternité, chômage, accident du travail, service militaire, guerre...)

Plafond de la sécurité sociale (PSS) : Le PSS sert de base au calcul des cotisations sociales. Dans le régime général des salariés, les cotisations de retraite de base sont principalement calculées sur les salaires ou revenus limités au plafond : on dit alors que les cotisations sont plafonnées. Ce plafond est actualisé en fonction de l'évolution du salaire moyen et il est revalorisé chaque année au 1er janvier.

Polycotisant : Personne qui a cotisé au régime général et au moins dans un autre régime de base.

Polypensionné : Personne ayant cotisé, durant sa vie professionnelle, à plusieurs régimes différents et bénéficiant, de ce fait, de retraites versées par plusieurs caisses au prorata de la durée passée dans chaque régime.

Régimes alignés : Régimes ayant choisi, en 1973, de se rapprocher du régime général en adoptant les mêmes règles pour le calcul des droits à retraite. Les régimes alignés regroupent le régime général des salariés, le régime des artisans et des commerçants et le régime agricole (pour les salariés agricoles).

Régimes non alignés : Régimes ayant leurs propres règles de calcul des droits à retraite, non alignées sur celles du régime général (exemple : le régime de la Société Nationale des Chemins de Fer, régime du personnel de la fonction publique...).

Régime complémentaire : Deuxième niveau de retraite obligatoire, complétant le régime de base (ex. régimes Arrco pour tous les salariés et Agirc pour les salariés cadres, régime Ircantec pour les agents non titulaires de l'Etat et des collectivités publiques, nouveau régime complémentaire obligatoire - NRCO - pour les commerçants...).

Retraite anticipée : Dans le régime général, l'âge légal de la retraite est fixé à 60 ans. Toutefois, depuis 2004, les assurés qui ont commencé à travailler avant 17 ans et qui respectent certaines conditions de durée d'assurance validée et de durée cotisée peuvent prendre leur retraite dès 56 ans (55 ans pour les assurés handicapés). On parle dans ce cas de « retraite anticipée ».

Salaire annuel moyen (SAM) : Salaire moyen servant de base au calcul de la pension de vieillesse. Il représente la moyenne des n meilleurs salaires, avec n compris entre 10 et 25, selon la génération de naissance de l'assuré.

Salaire moyen par tête (SMPT) : Il correspond à la masse salariale totale divisée par l'emploi total, et ajusté de l'inflation.

Surcote : Majoration appliquée au montant de la future pension d'un assuré âgé de 60 ans au moins et qui choisit de continuer à travailler, alors qu'il a atteint la durée d'assurance nécessaire pour bénéficier d'une retraite à taux plein. Le taux de surcote est compris entre 0,75% et 1,25% par trimestre supplémentaire, en fonction de l'âge de l'assuré et du nombre de trimestres de surcote dont il bénéficie.

Système par capitalisation : Mode d'organisation des systèmes de retraite dans lequel les cotisations d'un assuré sont placées à son nom durant sa vie active (placements financiers et immobiliers, dont le rendement varie en fonction des taux d'intérêt), avant de lui être restituées sous forme de rente après l'arrêt de son activité professionnelle. La constitution du capital peut s'effectuer à titre individuel ou dans un cadre collectif (accord d'entreprise). En France, seuls les systèmes de retraite dits sur-complémentaires (ex. : le PERP, ou plan d'épargne retraite populaire) fonctionnent selon le principe de la capitalisation.

Système par comptes notionnels : Dans les systèmes par comptes notionnels, chaque assuré est titulaire d'un compte individuel. Les cotisations acquittées chaque année par un assuré créditent « virtuellement » son compte et leur accumulation forme un « capital virtuel ». Celui-ci est revalorisé chaque année selon un index représentatif du taux de progression de l'assiette des cotisations du régime. Différentes assiettes peuvent être définies afin de couvrir des périodes particulières (chômage, éducation des enfants...). Le montant de la pension liquidée dépend du capital virtuel accumulé auquel est appliqué un coefficient de conversion dont la valeur dépend de l'âge effectif de départ à la retraite et de l'espérance de vie à cet âge. Ainsi, la pension n'est pas calculée en fonction du nombre d'années validées mais en fonction du nombre de points comptabilisés tout au long de la carrière. Les régimes par points sont des régimes par comptes notionnels.

Système par répartition : Mode d'organisation des systèmes de retraite fondé sur la solidarité entre générations. Les cotisations versées par les actifs au titre de l'assurance vieillesse servent immédiatement à payer les retraites. L'équilibre financier des systèmes de retraite par répartition est fonction du rapport entre le nombre de cotisants (population active, taux de croissance des revenus) et celui des retraités. Le système français de retraite est fondé sur le principe de la répartition.

Taux plein : Taux maximum de calcul d'une retraite, pour un assuré justifiant de la durée d'assurance nécessaire, tous régimes confondus. Pour le régime de base des salariés du privé, le taux plein est de 50%. Ce taux est appliqué au salaire annuel moyen. Peuvent aussi obtenir une retraite au taux plein, quelle que soit leur durée d'assurance, les personnes ayant atteint 65 ans et les personnes se trouvant dans une situation particulière (reconnues inaptes au travail, invalides, anciens combattants, anciens prisonniers de guerre, anciens déportés ou internés politiques, ...).

Trimestre assimilé : Trimestre n'ayant pas donné lieu à cotisations, mais néanmoins pris en compte, sous certaines conditions, pour le calcul de la durée d'assurance (ex. : périodes de chômage indemnisé).

Trimestre cotisé : Trimestre ayant donné lieu à versement de cotisations, calculées sur les salaires.

Trimestres validés : Ensemble des trimestres - cotisés, assimilés ou équivalents - pris en compte pour le calcul de la durée d'assurance.

Bibliographie

- Beaumel C., Vatan M. (2008), « La situation démographique en 2006 », Insee résultats n. 84, août 2008.
- Bonnet C., El Mekkaoui de Freitas N. (2004), « Retraite et patrimoine financier des ménages », Etudes et résultats n°330, août 2004.
- Caillot L., Chaput H., Colin C., El Mekkaoui de Freitas N., Michaudon H. (2003), « Echantillon interrégimes de cotisants (EIC). Tome 1 : Procédures statistiques de constitution de l'EIC » DREES, Série Statistique, working paper n° 50, mars 2003.
- Chambaz C., Hourriez J-M., Legendre N. (1999), " Le niveau de vie des retraités en 1995 ", Retraite et société n°28/1999.
- CNAV (2007), Recueil statistique 2007.
- Commission des Comptes de la Sécurité Sociale (CCSS) (2008), Rapports annuels sur la sécurité sociale de juin et septembre 2008 <http://www.securite-sociale.fr/chiffres/ccss/ccss.htm>.
- Conseil d'Orientation des Retraites, documents de travail www.cor-retraites.fr.
- Debrand T. et Privat A.-G. (2004), « Salaires individuels et évolutions macroéconomiques en France », Revue de l'OFCE, n° 89, avril, pp. 271-301.
- Insee (1999), « Le modèle de microsimulation dynamique DESTINIE », document de travail de la Direction des Études et Synthèses Économiques de l'INSEE, G 9913, août 1999.
- Mincer J. (1974), « Scholling, Experience and Earnings », Columbia University Press, New York.
- Mincer J., Polachek S. (1974), « Family Investments in Human Capital : Earnings of Women », Journal of Political Economy, vol 82, n° 2, 2^{ème} partie, pp. S76-S108.
- Ministère de la Défense, bilan social, www.defense.gouv.fr
- Poubelle V., Albert C., Beurnier P., Couhin J., Grave N. (2006), « Prisme, le modèle de la Cnav », Retraite et société, n°48, pp. 202-215
- Robert-Bobée I. (2006), « Projections de population 2005-2050, pour la France métropolitaine », Insee Résultats n° 57, septembre 2006.