

CONSEIL D'ORIENTATION DES RETRAITES
Séance plénière du 5 mars 2020 à 9h30
« Point sur les modèles de microsimulation »

Document N° 10
<i>Document de travail, n'engage pas le Conseil</i>

Le modèle Pablo de microsimulation de la fonction publique de l'État

Benoît Buisson, Service des retraites de l'État

CONSEIL D'ORIENTATION DES RETRAITES

Séance plénière du 5 mars 2020

« Modèles de microsimulation »

Le modèle Pablo de microsimulation de la fonction publique de l'État

Benoît Buisson, Service des retraites de l'État

Les modèles des régimes de retraite visent à projeter leur situation en termes démographique et financier à un horizon d'une cinquantaine d'années. Ils peuvent concerner l'ensemble des régimes ou un régime particulier. Pour le régime de la fonction publique de l'État, historiquement la direction du Budget assurait l'établissement des projections à long terme à l'aide du modèle Ariane. Ariane fonctionnait en regroupant la population par génération et en quelques grands groupes (civils, militaires, la Poste et Orange), afin de constituer des individus moyens par génération. Toutefois, cette méthode de simulation dite de *méso-simulation* ne permettait pas de prendre en compte toutes les finesses de la réglementation du régime de la fonction publique de l'État, ni d'étudier précisément les impacts de changements de réglementation.

En réponse au relevé d'observations de l'audit Ariane mené par la Cour des comptes en 2009, le service des retraites de l'État (SRE) et la direction du Budget ont suggéré la création d'un modèle de microsimulation de long terme. L'intérêt de ce modèle est de s'appuyer sur les données individuelles issues des comptes individuels retraite (CIR) qui retracent les droits de tous les agents ayant cotisé « un jour » au régime et pas uniquement ceux qui sont encore présents. Elles présentent une richesse importante en offrant une vision complète des droits acquis tout au long des années et pas seulement une vision instantanée de la situation de l'agent une année donnée. Dans ce contexte, et dans l'idée de fournir des projections précises sur des populations fines de fonctionnaires, il a été décidé d'engager la construction d'un modèle de microsimulation dynamique au SRE. En effet, même s'il est plus lourd à mettre en place et à maintenir, seul ce type de modèle permet de prendre en compte toute la finesse des règles de calcul des retraites, souvent non linéaires, les disparités des carrières entre agents, les changements de comportements attendus suite aux réformes.

Amorcé en 2012, le modèle Pablo est opérationnel depuis l'année 2017 et permet de répondre à trois grands types d'utilisations :

- effectuer des projections sur le régime de retraite actuellement en vigueur ;
- calculer les engagements de l'État en matière de retraite ;
- réaliser des évaluations ou études d'impact de politiques publiques (*ex ante* ou *ex post*).

Le premier objectif recouvre une évaluation financière de la charge et des produits du régime en fonction de différents scénarios démographiques et économiques. Cette projection financière permet une analyse des besoins de financement du régime à long terme. Au niveau macro-économique, l'utilisation du modèle Pablo permet de quantifier à long terme l'évolution de la pension moyenne ainsi que le nombre de cotisants et de retraités. Le modèle permet aussi de décrire les caractéristiques des nouveaux pensionnés une année donnée : âge à la liquidation, pension, durée acquise par exemple. Au-delà des évolutions moyennes, il s'agit d'analyser également la distribution des variables précédentes et son évolution au fil du temps, en distinguant des populations particulières. Dans ce cadre, le modèle Pablo est notamment utilisé pour répondre à l'exercice annuel de projection du COR de l'ensemble des régimes de retraite. Le modèle Pablo est également utilisé pour calculer les engagements de l'État en matière de retraite pour le régime de retraite des fonctionnaires civils et des militaires. Ces engagements, d'un montant avoisinant les 2 000 milliards d'euros, figurent en annexe du compte général de l'État. Il est d'usage également de calculer, en plus des engagements de retraite, le besoin de financement actualisé du régime ou dette implicite *ex ante*. Cet objectif de calcul des engagements a conditionné une partie de la structure du modèle.

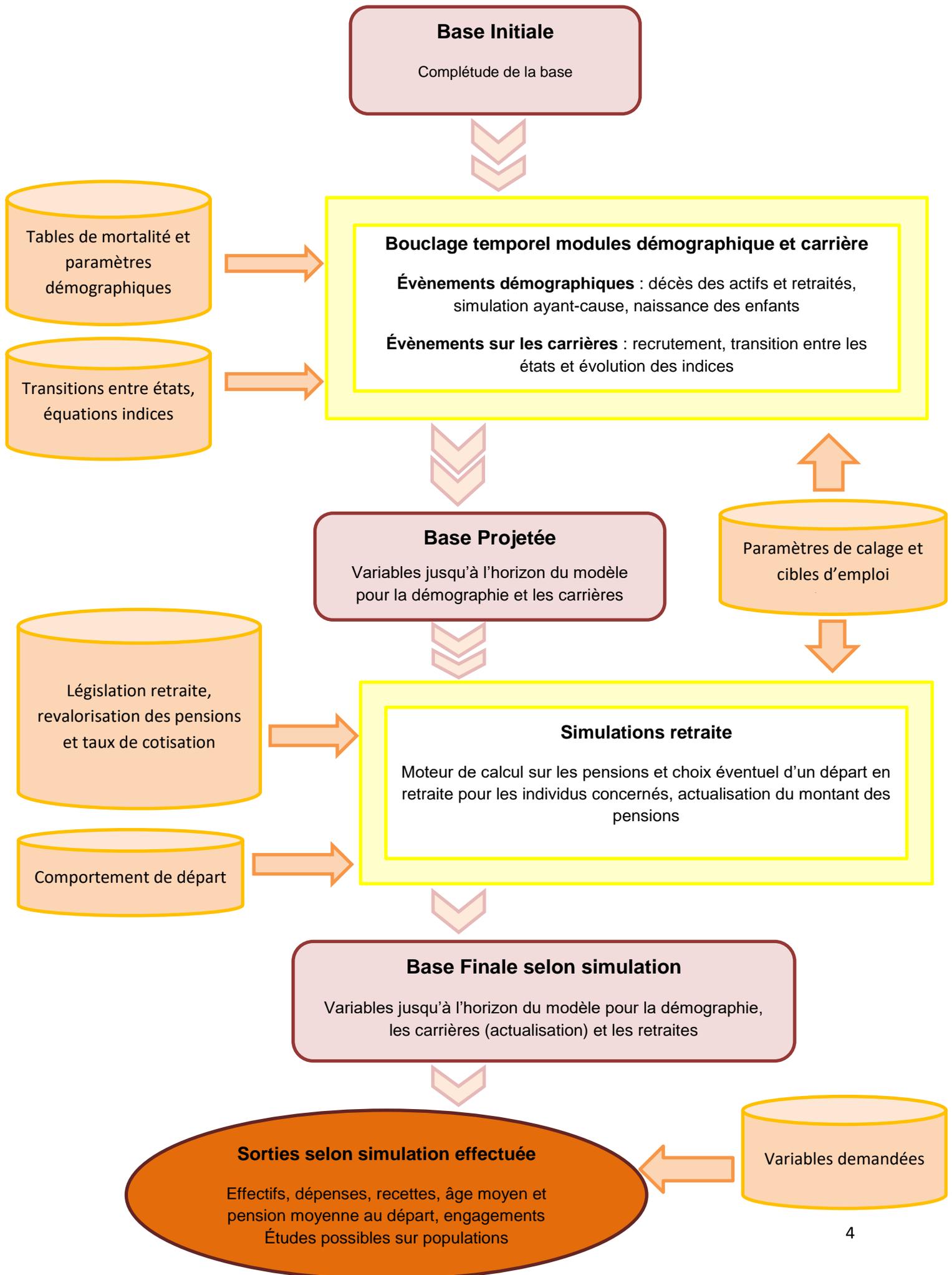
L'objectif de *réaliser des évaluations et des études d'impacts de politiques publiques* est particulièrement important. Dans le cas d'une évaluation *ex ante* (impact attendu d'une politique publique dans ce cas), la situation de référence pour le modèle est la législation en cours. Le modèle

Pablo permet de projeter les situations individuelles – en fonction de tout un jeu d’hypothèses et d’aléas – jusqu’à l’horizon du modèle. Il est aussi possible d’introduire une législation différente et de réaliser des projections pour déterminer l’impact d’une réforme donnée. Cette évaluation de l’impact potentiel est mesurée par toute une gamme de variables : montant moyen d’une pension, âge moyen à la liquidation, nombre d’années cotisées à la liquidation par exemple. Outre les niveaux agrégés, il est tout à fait possible, et souhaitable, d’étudier l’impact *ex ante* d’une politique publique par population. Cela permet par exemple de déterminer les *gagnants* et les *perdants* d’une réforme donnée, de faire du ciblage des effets d’une réforme sur des populations particulières. Il est aussi possible de faire différents scénarios sur le jeu d’hypothèses – notamment adaptation ou non des comportements - pour donner une amplitude de l’impact, en mettant en évidence la sensibilité aux hypothèses retenues. Dans le cas d’une évaluation *ex post*, le modèle Pablo permet de simuler des départs en retraite sur la base d’une législation qui n’est plus en vigueur. La comparaison avec les flux constatés permet de cerner l’impact d’une réforme *ex post*, éventuellement d’étudier les changements de comportements induits par la réforme.

I) Architecture et sources statistiques du modèle

L’architecture du modèle est exposée de manière synthétique dans le schéma page suivante. Les principes qui régissent cette architecture sont relativement simples. Dans un premier temps une base initiale est constituée à partir des informations présentes d’une part dans le compte individuel retraite et d’autre part dans la base des pensions. Le compte individuel retraite permet de retracer toute la carrière d’un agent, année par année (position statutaire, quotité de temps travail, type de temps partiel, modalité de congés...) en fournissant également des éléments de rémunération, uniquement toutefois sur le traitement indiciaire (pas d’information sur les primes). La base des pensions présente quant à elle des éléments détaillés du calcul de la pension à la liquidation comme la durée de service acquise, les bonifications, l’indice à la liquidation ou encore le taux de décote ou de surcote. Cette base initiale donne ainsi une vue complète du régime à la date de référence. Dans un second temps, toute une série d’évènements va être simulée pour passer de cette base initiale à la base projetée, dont l’horizon sera 2070. Ces évènements sont soit de nature démographique, décès – mariage – naissance, soit de nature économique avec la projection des carrières : situation sur le marché du travail et quotité de travail, indice, promotion éventuelle, recrutement dans la fonction publique de l’État. Ces évènements seront générés volontairement jusqu’à un âge « avancé ».

À ce stade, aucun comportement de départ en retraite n’est simulé. Cette *base projetée* sera générée un nombre limité de fois et stockée dans l’objectif de pouvoir donner lieu à des simulations multiples en termes de retraite. L’utilisateur pourra ensuite demander une simulation pour, par exemple, disposer d’une vue du régime dans la législation existante à long terme, ou encore effectuer une variante de politique publique en la matière. Pour cela l’utilisateur spécifiera principalement à la fois la législation retraite sur laquelle il veut travailler et les comportements de départ à la retraite des individus de la fonction publique de l’État (taux plein par exemple). Par simulation, l’utilisateur obtiendra ainsi une base *finale* qui contient l’ensemble des évènements démographiques sur la carrière et la retraite. À partir de cette base finale, il sera possible d’obtenir des indicateurs synthétiques comme la pension moyenne, les effectifs de retraités par année ou encore la masse des pensions à verser. Il pourra également réaliser des analyses longitudinales par population ou par génération pour aborder des problématiques comme les mécanismes de redistribution au sein d’une génération ou l’équité intergénérationnelle. La phase *simulation retraite* pourra être répétée sur la même *base projetée* selon les demandes et les études à réaliser.



Le champ du modèle Pablo est l'ensemble des affiliés au régime de la fonction publique de l'État, qu'ils soient civils ou militaires. Pour constituer la *base initiale* du modèle il a été nécessaire de déterminer ce champ à une date de référence : le 31 décembre 2015. Le champ de la base initiale est séparé en deux populations : les *retraités* à la date de référence et les *affiliés non retraités* à cette même date.

Pour les retraités à la date de référence, les sources traditionnelles du SRE sur les pensions sont directement utilisées. Les différentes informations collectées sont de trois types : renseignement générique sur le pensionné (date de naissance, sexe...), informations relatives à la liquidation de la pension (durée dans le régime, durée autres régimes, taux de décote...) et celles qui concernent le montant de la pension versée jusqu'à la date de référence. Initialement il était prévu de faire un échantillonnage au vingtième de la population des retraités. Suite à une préconisation de la Cour des comptes dans le cadre du calcul des engagements en termes de retraite, la population est restée exhaustive pour les retraités à la date de référence (2 millions d'individu environ).

Pour les affiliés non-retraités à la date de référence, la source d'information principale est constituée par les informations issues du compte individuel retraite (CIR). Cette nouvelle source était à l'origine même du modèle et explique également le choix de la date de référence au 31 décembre 2015. Sur cette population, le champ correspond donc à l'ensemble des affiliés qui disposent de droits dans le régime, qui ne sont pas en retraite à la date de référence et qui sont donc susceptibles de partir à la retraite après la date de référence. Le champ est donc ici plus large que les cotisants à la date de référence. Il comprend notamment des personnes en disponibilité ou qui ont été radiés des cadres mais qui ont acquis des droits dans le régime par le passé.

Pour toutes les personnes qui appartiennent au champ un certain nombre de caractéristiques sont collectées grâce au CIR : sexe, date de naissance, individu civil ou militaire, administration d'origine, emploi de type sédentaire ou actif par exemple. Un tirage stratifié au vingtième est ensuite effectué pour déterminer l'échantillon de référence du modèle. La variable de stratification distingue la population selon le caractère civil / militaire, l'administration d'origine, la nature de l'emploi occupé (actif ou sédentaire). L'échantillon est ainsi représentatif sur ces 3 variables. Puis dans chaque strate la population est triée par génération et sexe, un individu sur 20 est sélectionné par strate, ce qui assure également une représentativité de l'échantillon sur ces deux autres variables. Une fois l'échantillonnage effectué, des caractéristiques supplémentaires, notamment de carrière, vont être collectées ou éventuellement imputées uniquement sur les individus retenus, soit un individu sur 20 (environ 200 000 individus). Ces informations sont principalement de deux types : un résumé de la carrière notamment sous forme de durées acquises au sein de la FPE au 31 décembre 2003, la chronologie de la carrière par trimestre entre le premier trimestre 2004 et le quatrième trimestre 2015 (état sur le marché du travail, quotité de travail, indice détenu chaque trimestre notamment). Pour cela, malgré l'importante montée en qualité des CIR, une phase de complétude des données est nécessaire pour obtenir des informations complètes pour chaque individu de l'échantillon. Cette complétude des données est opérée par imputation : chaque individu avec des données incomplètes de carrière se voit attribuer les données d'un individu donneur de l'échantillon, relativement proche de lui en termes de caractéristiques (administration, grade, âge d'entrée dans la fonction publique...). Une fois cette opération de complétude terminée, tous les individus échantillonnés disposent de l'ensemble de leur chronologie de carrière jusqu'à la date de référence. Ils pourront ainsi, via les différents événements générés par les modules, être projetés jusqu'à un horizon lointain. Ces événements seront regroupés en trois modules : démographie, carrière et retraite.

II) Description des modules démographie, carrière et retraite

Trois événements de nature démographique sont simulés dans le modèle Pablo : mariage, naissance et surtout décès. Le mariage et les naissances consistent « simplement », pour chaque affilié du régime, à lui imputer un conjoint marié éventuel (avec la date de naissance de celui-ci) et des enfants avec date(s) de naissance associée(s). Ces événements sont importants dans le cadre des droits familiaux – bonification ou majoration de durée – et bien sûr dans le cas des pensions de réversion. Que cela soit pour les mariages et les enfants, ces événements sont calés sur les tables de nuptialité et de fécondité de l’Insee, en supposant que les fonctionnaires n’ont pas un comportement particulier en la matière.

Pour établir des projections à long terme sur les régimes de retraite, il s’avère crucial de bien simuler la mortalité par âge et sexe jusqu’à un horizon temporel lointain. La simulation des décès se fait en appliquant des quotients de mortalité projetés. Chaque individu, en fonction de ses caractéristiques notamment son sexe et son âge, se verra affecter une probabilité de décéder chaque année. Suite au tirage d’une variable aléatoire individuelle, le modèle simulera ou non le décès de l’individu une année donnée. Si le décès est simulé l’année en question, une date de décès sera sélectionnée dans l’année en respectant la saisonnalité actuelle des décès.

Appliquer des quotients de mortalité qui ne seraient pas en adéquation avec la mortalité dans le régime pourrait entraîner des erreurs de projections importantes sur le nombre de cotisants, de retraités ou encore sur la masse des pensions à verser. Ces imprécisions de projections seraient d’autant plus préjudiciables que les réformes, depuis 2003, ont placé au centre de l’analyse du système de retraite français des indicateurs comme l’espérance de vie par génération, ou encore des ratios de type durée de retraite sur l’espérance de vie ou durée de carrière sur durée de retraite. De nombreuses études¹ ont explicité les origines des différences individuelles d’espérance de vie, différences individuelles qui expliquent que les personnes de la fonction publique de l’État (FPE) ont une espérance de vie plus élevée que la moyenne de la population française. Dans ce contexte il ne s’avère pas pertinent d’utiliser directement les quotients de mortalité projetés de l’Insee. La voie choisie pour simuler les décès dans le modèle Pablo est celle des modèles relationnels dits de Hannertz. Le principe consiste à établir une relation entre les quotients de mortalité de la population française dans son ensemble et ceux des populations de la FPE en introduisant l’âge en paramètre supplémentaire. Cela permet de prendre en compte le fait que l’écart de mortalité se modifie selon l’âge. Une fois cette relation établie par population, elle est maintenue constante sur l’horizon de projection. Cela permet d’obtenir des tables de mortalité, par âge, jusqu’en 2070 pour chacune des 23 populations de la FPE mises en évidence. Une méthode spécifique est employée pour estimer les quotients de mortalité aux grands âges.

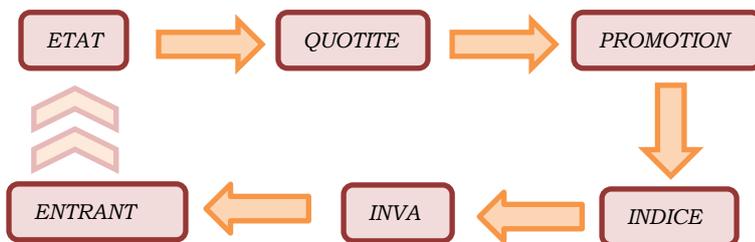
La présence éventuelle d’un conjoint marié ainsi que d’enfant(s) ainsi que la simulation d’une date de décès vont permettre la simulation des pensions de réversion. Dès que le décès d’un individu de type ayant-droit va être simulé, on regarde la présence d’un conjoint marié vivant à la date du décès ou éventuellement d’enfants de moins de 21 ans en vie (pour les pensions de type orphelin). Si tel est le cas, les individus en question vont faire leur entrée dans la base des affiliés au régime.

Les trois événements principaux du module carrière sont simulés sur un rythme trimestriel : détermination de l’état sur le marché du travail (ETAT), détermination de la quotité de travail si

¹ Voir par exemple COR (2014) : « Analyse de quelques disparités d’espérance de vie », groupe de travail du 19 mars 2014.

l'individu est en activité au sein de la FPE (QUOTITE) et enfin détermination de l'indice majoré chaque trimestre (INDICE).

Trois évènements du module carrière sont d'une périodicité annuelle, en se déroulant le premier trimestre de chaque année : changement de catégorie hiérarchique (PROMOTION), passage en invalidité (INVA) et enfin entrée au sein du régime de la FPE (ENTRANT). Ainsi le premier trimestre de chaque année de simulation les évènements vont s'enchaîner de la façon suivante :

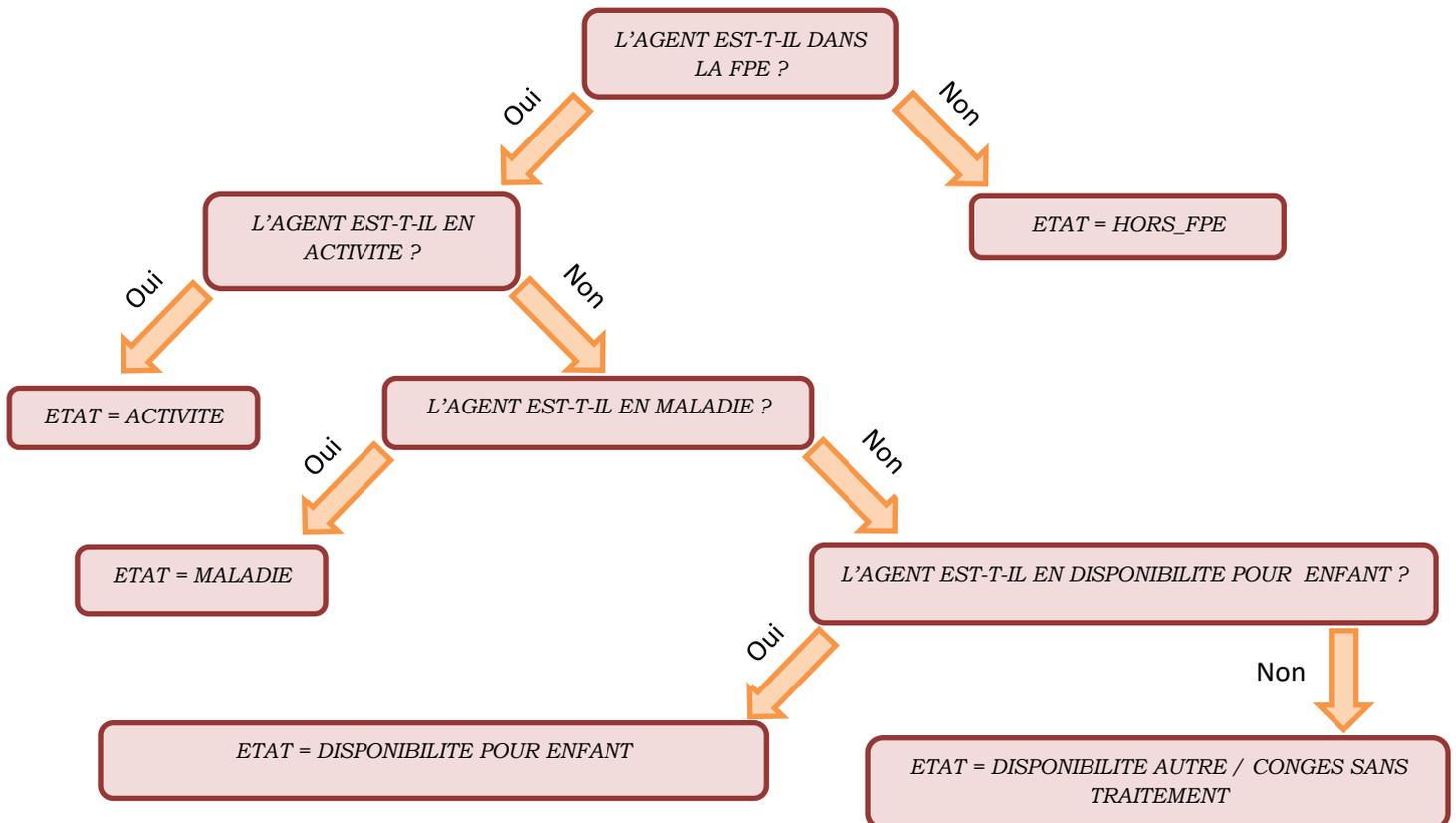


Suite à l'évènement ENTRANT, la séquence des évènements du premier trimestre sera terminée. Les évènements du deuxième trimestre vont pouvoir se succéder : ETAT, QUOTITE et INDICE et ainsi de suite.

Le principe général retenu pour la simulation des évènements du module carrière consiste à modéliser les comportements observés sur le passé (état sur le marché du travail, indice...) et de reproduire ces comportements passés jusqu'à un horizon de projection lointain. Il n'est donc pas introduit des scénarios de changement ou de rupture de comportements à ce stade. Les comportements sont modélisés à partir de la base de données constituée grâce au CIR et via des techniques économétriques classiques. L'évènement ETAT consiste à déterminer chaque trimestre la situation sur le marché du travail parmi les possibilités suivantes : activité (au sein de la FPE), maladie (au sein de la FPE), en disponibilité pour enfant, en disponibilité pour autres raisons ou congés sans traitement, hors de la FPE. Ces modalités ont été choisies en lien avec le calcul des durées, durée de service notamment, au sein de la FPE au moment du passage à la retraite. Ainsi pour les états « activité » et « maladie » la durée de service s'incrémentera en fonction de la quotité de travail et de la sur-cotisation éventuelle, cette durée s'incrémentera de 90 jours par trimestre dans le cas de disponibilité pour raisons familiales et ne s'incrémentera pas pour les deux derniers états (autres disponibilité ou congé sans traitement, hors de la FPE). Le principe générique de la modélisation de la situation sur le marché du travail consiste, par population, à établir des modèles logistiques dichotomiques qui vont permettre de calculer des probabilités de transition d'un état à l'autre en fonction des caractéristiques individuelles de la personne. Les modèles logistiques permettent de mesurer l'association entre la survenue d'un évènement (variable expliquée) et les facteurs susceptibles de l'influencer (variables explicatives). La variable expliquée peut être binaire (c'est-à-dire de type « oui / non » ou « vrai / faux » pour les modèles logistiques dichotomiques) ou posséder plus de 2 modalités (modèles logistiques polytomiques). Le choix des variables explicatives intégrées au modèle logistique est basé sur une connaissance préalable des déterminants des carrières individuelles. Les variables explicatives qui seront introduites dans le *modèle peuvent être* quantitatives ou qualitatives. Lors de la simulation de chaque évènement, une probabilité à chaque période sera calculée en fonction des caractéristiques individuelles de chaque agent (par exemple la probabilité de changer d'indice le trimestre en question). Suite à un tirage d'un nombre aléatoire compris entre 0 et 1, la comparaison de cette probabilité avec ce nombre aléatoire va déclencher l'évènement ou non. Par exemple si chaque agent avait une

probabilité de changer d'indice de 30 %, le tirage d'un nombre aléatoire et sa comparaison avec la probabilité assure que 30 % des agents vont voir leur indice progresser et 70 % stagner.

La modélisation de l'évènement se fera selon l'arbre de décision suivante.



Pour les individus qui sont en activité ou en maladie le trimestre considéré, la deuxième étape consiste à modéliser la quotité de travail. Cet évènement QUOTITE sera modélisé en deux temps. Dans un premier temps, *via* un modèle logistique classique, on simule l'évènement dichotomique « changement ou non de la quotité de travail » entre le trimestre en cours et le trimestre précédent. Si la simulation aboutit à un changement de quotité le deuxième temps consistera à déterminer cette nouvelle quotité. *Via* une modélisation logistique polytomique une quotité de travail sera déterminée parmi les possibilités suivantes : 50 %, 60 %, 70 %, 80 %, 90 % et 100 %. Les agents de retour de disponibilité, ou d'un état « hors-fpe », passeront uniquement dans cette seconde phase car pour eux le problème du changement de quotité par rapport à la période précédente ne se pose pas. Enfin les entrants se verront affecter une quotité de travail de 100 % pendant leur trimestre d'arrivée dans la FPE.

L'évènement PROMOTION consiste à modéliser le passage d'une catégorie « statutaire » à une autre : soit un passage de la catégorie C à la catégorie B ou un passage de la catégorie B à la catégorie A. Cette modélisation s'effectue *via* une régression logistique classique par population issue d'un croisement du sexe avec la tranche d'âge et la catégorie « statutaire » le trimestre précédent. Cet évènement est annuel.

Actuellement dans le modèle Pablo, la simulation du salaire consiste uniquement dans la FPE en la simulation de l'indice majoré chaque trimestre, en lien avec la législation actuelle en matière de

retraite dans la fonction publique où les cotisations et les droits sont calculés en référence au seul traitement indiciaire. Il est prévu, dans une deuxième version du modèle, d'intégrer une modélisation des primes et d'arriver à une simulation à chaque période de l'ensemble des composantes du salaire.

Pour les individus en activité et en maladie, l'évènement INDICE se déroule en deux temps : changement d'indice ou non le trimestre considéré, et si l'indice est modifié, détermination du nouvel indice majoré dans un deuxième temps. La première étape est réalisée grâce à un modèle logistique par population croisant le sexe, la tranche d'âge et l'état (activité ou maladie) le trimestre précédent. En cas de changement d'indice, la simulation du nouvel indice – par tranche d'âge – se fait via un modèle linéaire. Les variables retenues sont, par exemple, le nombre d'années de carrière, l'âge de début de carrière, le fait d'être actif ou sédentaire, la catégorie hiérarchique ou encore le ministère d'appartenance. Le principe de la simulation de l'indice majoré consiste donc, comme pour les autres évènements, à reproduire les évolutions passées jusqu'à un horizon de projection lointain. Le choix a été fait de ne pas introduire de grille par corps et de ne pas projeter l'évolution de l'indice selon ces grilles. Lorsque la réforme sur les parcours professionnels, la carrière et la rémunération (PPCR) est apparue, après la mise en place du modèle Pablo, il a été nécessaire d'introduire des modifications dans la modélisation. En effet cette réforme a introduit une rupture, surtout dans le court terme, des évolutions d'indice suite à la mise en place de nouvelles grilles. Un effet de cette réforme a donc été calculé par la direction du Budget et introduit dans le modèle, via un processus de calage.

Dans le modèle Pablo, le passage en invalidité est un évènement du module carrière (évènement INVA). Suite à ce passage en invalidité une pension est calculée dans le modèle retraite. Lors de l'évènement annuel INVA, une probabilité de départ en invalidité sera ainsi calculée pour chaque agent. Cette probabilité de mise en invalidité dépendra de caractéristiques comme l'âge, le nombre de trimestres passé en état « maladie » ou encore le ministère employeur. Suite au tirage d'un nombre aléatoire, si l'agent devient invalide cela signifie qu'il part à la retraite et donc que les évènements du module carrière vont cesser pour l'agent en question.

Le dernier évènement du module carrière est l'évènement ENTRANT. Il s'agit de simuler chaque année un certain nombre de nouvelles affiliations dans le régime. En fait cet évènement va être simulé en deux temps : lors du module carrière d'une part, lors du module retraite d'autre part. Pour connaître le nombre d'entrants dans le régime une année donnée il est nécessaire de disposer d'effectifs cibles et du nombre de sortants (décédés en activité et surtout retraités). Les effectifs cibles sont fournis, par année et par grand ministère employeur, par la direction du Budget en fonction des hypothèses retenues de variation de l'emploi public notamment dans le cadre du projet de loi de finances. Le nombre de départs à la retraite n'est pas connu lors du module carrière. Il est néanmoins indispensable de générer des nouveaux entrants à ce stade, afin de simuler des carrières individuelles à partir de l'année d'entrée. Dans ce cadre, il va être estimé – dès le module carrière - un nombre de retraités par année de manière très simple, plutôt dans l'idée de maximiser un nombre d'entrants. Lors du module retraite le nombre de retraités sera effectivement connu précisément, de ce fait certains entrants pourront ensuite être supprimés afin de respecter par année l'effectif cible.

Les évènements du module carrière vont donc s'enchaîner entre la date de début de simulation et la date de fin, généralement 2070. Pour chaque individu, en attente de connaître la date de départ à la retraite, les évènements associés à la carrière vont être simulés jusqu'à un âge avancé (75 ans par exemple) avec une modélisation spécifique pour la fin de carrière. Une fois que les évènements de nature démographique et de type carrière ont été simulés jusqu'en 2070, on obtient une base projetée avec l'ensemble des évènements simulés, base qui sera en entrée du module retraite.

Toute simulation sur les retraites se caractérisera par une législation donnée en matière de retraite et par le choix d'un comportement de départ à la retraite (départ au taux plein par exemple). Ainsi à une date t donnée, le modèle simule l'éventuel départ à la retraite des individus en fonction d'une législation et d'hypothèses de comportement de départ.

Au sein du modèle retraite les deux premières opérations consisteront d'une part à calculer pour chaque agent le couple de variables (date d'ouverture des droits, limite d'âge) et d'autre part à calculer toutes les durées nécessaires (durée de services, durée d'assurance, bonifications ...) pour la législation demandée, uniquement à la date d'ouverture des droits de chaque agent. Ces phases sont mises en amont du cœur du module retraite car elles sont chronophages en temps de calcul. Suite à ces opérations des bases peuvent être stockées pour un certain nombre de législation, dont la législation actuelle, avec les durées à la date d'ouverture des droits pour chaque agent. Cela permettra aux autres opérations du module retraite de s'enchaîner plus rapidement, notamment la mise à jour des durées entre la date d'ouverture des droits et la date de départ à la retraite, et d'être plus réactif en cas de simulations répétées. Le module retraite concerne les individus de type ayant-droit (droit direct) et de type ayant-cause (droit dérivé). Le déroulement du module retraite pour les droits directs est décrit dans le bloc ci-dessous.

Bloc sur la simulation des droits directs

Constitution d'un fichier comprenant les individus concernés par ce bloc, affiliés non retraités à la date de début de simulation.

Calcul de la Date d'Ouverture des Droits et de la limite d'âge si la législation demandée impose des changements en la matière par rapport à la législation actuelle.

Départ à la retraite des agents qui ont dépassé leur limite d'âge à la date de simulation.

Boucle mensuelle entre janvier de l'année de début de simulation et décembre de l'année de fin de simulation :

- Sélection des individus concernés par un départ potentiel sur la période ;
- Pour chaque individu concerné par un départ potentiel
 - o Mise à jour des durées pour la période ;
 - o Décision ou non de liquidation ;
 - o Calcul du montant de la pension à la liquidation pour les agents qui liquident ;
 - o Cas particulier des départs pour invalidité : liquidation et calcul du montant de pension ;
 - o Constitution d'une base des personnes qui ont liquidé leur retraite.

Fin de la boucle temporelle

Boucle temporelle sur la revalorisation des pensions de droit direct

Boucle temporelle sur le calcul des cotisations salariales

Il est à noter que pour un certain nombre de projections, notamment dans le cadre des projections du COR, on impose au modèle Pablo d'être calé sur les prévisions de court terme (à horizon de 5 ans) issues du modèle Paco. Dans ce cas, avant la boucle mensuelle mentionnée dans le schéma ci-dessus, il est procédé à un calage sur le modèle de court terme. Celui revient à définir, en lien avec le modèle de comportement de départ choisi, une date de liquidation pour tous les agents concernés avant d'effectuer le calcul de la pension dans la boucle temporelle. Suite au calage sur les prévisions de court terme, basées sur les comportements de départs à la retraite les plus récents, tous les individus ne partent plus au taux plein (comportement de départ à la retraite de référence du module Pablo).

Enfin une partie du module retraite concerne uniquement la réversion et les individus de type ayant-cause. Il est utile de rappeler ici que la date d'arrivée de ces individus dans le régime a été déterminée dans le module démographie. Le module retraite se contente donc de calculer la pension à la liquidation, suite au décès de l'ayant-droit, et de revaloriser cette pension jusqu'au décès de l'ayant-cause concerné.

Après le passage du module retraite toutes les informations du modèle de microsimulation sont disponibles pour être analysées. Quand les données en sortie de modèle sont demandées selon un format précis de restitution, pour l'exercice de projection du COR notamment, des programmes sont regroupés dans un module « sortie » afin de confectionner les tableaux au format demandé.

III) Test, validation et utilisations

Tout au long de la construction du modèle de nombreux tests ont été effectués. Le premier objectif était, en partant de la situation à la date de référence, de « reconstituer le passé » en simulant les événements à rebours. Toutes les relations économétriques mises en évidence ont été testées de cette manière afin d'éviter les calages qui pourraient engendrer des difficultés d'interprétation en projection.

Ensuite des premières simulations ont été effectuées avec la législation actuelle jusqu'à un horizon de 2070. Les résultats de ces premières simulations ont été finement analysés pour juger de leur pertinence. Des tests plus ponctuels ont été effectués afin de cibler l'analyse sur une composante précise du modèle.

Par exemple, des tests ont été effectués sur l'impact du choix d'une modélisation spécifique pour la simulation des décès en se ciblant uniquement sur la population des ayants-droits présents dans le régime au 31 décembre 2012. L'hypothèse alternative ici était de prendre directement les quotients de mortalité projetés, calculés par l'Insee, en supposant donc que la mortalité était identique entre les individus de la FPE et la population française dans son ensemble. Cette étude, menée en 2012, montrait que l'écart annuel en termes de nombre de décès simulé atteignait son maximum en 2019 avec 6 800 décès et l'écart cumulé en 2029 avec plus de 60 000 décès. Simuler les décès dans la FPE avec une modélisation propre à la fonction publique engendre donc 60 000 décès de moins à horizon 2029, par rapport à l'utilisation classique des quotients de mortalité projetés de l'Insee. Ces différences constatées sur le calendrier des décès se répercutent bien évidemment sur le montant des pensions à verser chaque année. L'écart entre le scénario FPE et le scénario Insee, atteindrait son maximum en 2034 avec plus de 2,2 milliards d'euros de montant de pensions. Cette mesure d'impact budgétaire annuel a été complétée par une analyse sur le cycle de vie des affiliés de la génération 1947. Le ratio « durée moyenne de retraite rapportée à la durée moyenne de vie » variait ainsi, pour les hommes

sédentaires de catégorie A de cette génération, de 27,5 % selon la mortalité Insee à 29,7 % avec la modélisation FPE.

Enfin le calcul des engagements de l'État en matière de retraite au 31 décembre 2015 a été effectué d'une part avec le modèle Ariane et d'autre part avec le modèle Pablo. Cette comparaison a été suivie attentivement par la Cour des comptes. De manière plus globale, cette dernière a mené un audit complet du modèle Pablo dans le cadre de son utilisation pour le calcul des engagements de l'État. Le modèle Pablo a ainsi été certifié par la Cour des comptes en décembre 2017.

Les utilisations régulières du modèle, notamment dans le cadre de l'exercice de projection du COR et du calcul des engagements de l'État, permettent des améliorations en continu de celui-ci. Dans le cadre de la réforme avec mise en place d'un système universel de retraite, le SRE a entamé la construction d'un nouveau modèle qui intègre le calcul des droits avec une législation en points (voir dernière partie). Les nouveaux développements portent désormais sur ce modèle, préfigurateur d'une version 2 de Pablo.

IV) Intérêts et cas d'usage

La construction d'un modèle de microsimulation est une opération coûteuse et complexe. D'un point de vue technique, elle nécessite de manipuler de grandes masses de données et se traduit souvent par des temps de traitement potentiellement très importants. Une autre difficulté de l'usage des modèles de microsimulation réside dans la génération de multiples événements qui introduit de nombreux effets aléatoires. La génération successive de l'ensemble des événements donnera ainsi des bases en sortie qui seront de fait différentes. En pratique, on ne mesure pas clairement et directement l'impact de cet aléa en raison des temps de traitement qui empêchent de recourir à des techniques de type *bootstrap* consistant à générer plus de bases projetées et à mesurer l'écart à la moyenne observée des résultats.

Dans ce contexte, quel est l'intérêt d'élaborer et d'utiliser un modèle de microsimulation ? En premier lieu, si l'objectif est d'analyser un impact de politique public à un niveau fin, sur de nombreux critères ou en regard de la distribution des effets, et non uniquement de la moyenne, la seule possibilité d'étude est bien l'utilisation d'un modèle de microsimulation. En second lieu, même lorsqu'on peut se limiter à l'analyse d'un effet moyen, l'utilisation d'un modèle agrégé revient à privilégier la notion d'agent représentatif. De fait, le premier article qui prônait l'utilisation de la microsimulation (Orcutt, 1957) était également une critique de la notion d'agent représentatif.

Or, la législation actuelle en termes de retraite rend particulièrement délicate l'utilisation de la notion d'agent représentatif. Les exemples sont nombreux en la matière et assez pédagogiques.

Par exemple, le calcul de la décote ou de la surcote fait intervenir un minimum entre deux *distances*. Relativement à la décote dans la fonction publique de l'État, il est important de distinguer d'une part les civils et les militaires qui ont des législations différentes et d'autre part, au sein des civils, les agents de la catégorie active des sédentaires qui n'ont pas les mêmes paramètres de calcul (âge d'annulation de la décote par exemple). Pour les civils sédentaires, le calcul de la décote dépend de la génération de l'agent, notamment via la durée de référence pour atteindre le taux plein. Avec une approche en termes d'agent représentatif, cet agent moyen représenterait les valeurs moyennes de

chaque génération de civils sédentaires. Ainsi, l'agent représentatif pour la génération 1955 des civils sédentaires ayant liquidé sa pension dans la fonction publique de l'État en 2019 serait parti à 63 ans et 11 mois avec une durée d'assurance de 175 trimestres (173 pour la durée d'assurance prise en compte pour la surcote) et une majoration de 4,2 % de la pension liée à la surcote, qui sont les valeurs moyennes constatées pour cette génération cette année-là. Hors ces données moyennes... ne sont pas cohérentes. Une personne de la génération 1955 qui part à 63 ans et 11 mois avec 173 trimestres devrait enregistrer une surcote de 8,75 % et non de 4,2 %. Ce point illustre qu'on ne peut pas raisonner avec un modèle d'agent représentatif avec la législation retraite actuelle en raison des mécanismes non linéaires (écrêtement des durées, calculs de la décote et de la surcote, prise en compte d'un éventuel minimum garanti par exemple) qu'elle contient. L'introduction d'un minimum garanti, qui peut se substituer au calcul classique de la pension lorsque son montant est plus élevé, est un autre exemple de la nécessité d'utiliser la microsimulation. Le minimum de pensions introduit en effet également des non-linéarités qui rendent délicates l'utilisation d'un agent représentatif.

L'exemple précédent montre qu'il faudrait multiplier les critères, par exemple par durée d'assurance en trimestres ou par âge de début de carrière, afin de définir des populations qui permettent de raisonner par agent représentatif. La multiplicité de ces critères à prendre en compte entraînerait la définition d'un grand nombre de populations qui rendrait un modèle semi-agrégé par population quasiment aussi complexe, et sans doute plus difficile à calibrer, qu'un modèle de microsimulation. Une voie possible de rapprochement entre un agent représentatif et les modèles de microsimulation peut éventuellement être trouvée en appliquant le modèle non pas à une multitude d'agents mais à un nombre plus réduit dit « agent type ». Cet agent type pourrait être un individu moyen selon certains critères ou un agent qui déroule une carrière type dans un certain grade. Il est donc en partie « fictif », ce qui de plus pourra engendrer des problèmes de représentativité, mais l'application d'un modèle à ce type d'agent peut avoir notamment une vertu pédagogique.

Par rapport à un modèle semi-agrégé statique, l'avantage d'un modèle de microsimulation dynamique est de bien prendre en compte les nombreux effets de structure internes à la population de la fonction publique de l'État. Ainsi la hausse tendancielle de la catégorie A, des âges de début de carrière de plus en plus tardifs, la fin programmée des instituteurs actifs de catégorie B sont autant de faits qu'il est primordial de prendre en compte dans les exercices de simulations.

Lors de l'exercice de comparaison du calcul des engagements de l'État en matière de retraites, entre les modèles Ariane et Pablo, les effets de structure ont expliqué une partie des différences entre les deux approches. Cet exercice de comparaison, lors du calcul des engagements au 31 décembre 2015, a été riche d'enseignements sur les apports d'un modèle de microsimulation. Le montant d'engagement calculé avec le modèle Pablo s'est révélé être supérieur d'environ 310 milliards d'euros à celui calculé avec le modèle agrégé Ariane, 1 845 milliards d'euros constants 2015 contre 1 535 milliards d'euros.

De multiples variantes de calcul, à partir du modèle Pablo, ont été menées afin de proposer une décomposition de cet écart. Ces variantes ont pu être réalisées uniquement avec Pablo, ce qui montre également la plus grande « souplesse » des modèles de microsimulation par rapport aux modèles agrégés. Sur les 310 milliards d'écart constaté, 46 provenaient d'une différence de champ. L'utilisation des comptes individuels retraite comme source du modèle Pablo, par rapport au Système d'Information sur l'Administration et la Sphère Publique (SIASP) pour le modèle Ariane, permet

d'être parfaitement en phase avec la population couverte par le régime. La source SIASP prend en compte les cotisants au 31 décembre une année donnée. Ainsi une personne en disponibilité au 31 décembre 2015, après avoir travaillé 10 ans dans la fonction publique, est bien couverte par la source CIR mais pas par le SIASP.

La deuxième source d'écart, déjà décrite, concernait l'utilisation de tables de mortalité spécifiques pour le régime et non des tables génériques Insee. L'écart généré sur les engagements d'utilisations de tables différentes de mortalité a été chiffré à 45 milliards.

L'essentiel de l'écart, 219 milliards, concernait le cœur des apports de la microsimulation. Les différences de pensions moyennes sont apparues comme la plus grande source d'écart avec un différentiel de 140 milliards. Cet aspect est profondément lié au thème précédemment traité sur les nombreuses non-linéarités qui interviennent dans le mode de calcul d'une pension. Raisonner sur modèles agrégés ou semi-agrégés, avec la définition d'individus moyens et le calcul de pension moyenne, génère de nombreuses incohérences. Ainsi le calcul d'une pension à la liquidation, via un moteur de calcul adapté, à un niveau individuel puis l'agrégation de ces montants de pension pour aboutir à une population moyenne par population donne des résultats forts différents de la définition d'un individu « moyen » avec une pension moyenne associée.

Enfin les 79 milliards d'euros d'écarts qui restent provenaient directement de l'apport de l'utilisation d'un modèle de microsimulation dynamique. Le modèle Ariane ne prenait pas bien en compte les déformations de structure de la population au fil des générations, par exemple via une hausse des départs de catégorie A notamment pour les femmes et de manière générale des carrières féminines plus complètes au fil des générations. Le modèle Ariane avait tendance à être trop calé sur les derniers flux de départs à la retraite sans prendre en compte la déformation des populations au fil du temps, bien appréhendée avec la source individuelle des CIR qui permet de connaître la carrière de tous les agents depuis leur entrée dans la fonction publique d'État. Cet exercice de comparaison sur le calcul des engagements a donc été particulièrement riche sur les apports de l'utilisation d'un modèle de microsimulation en lieu et place d'un modèle agrégé.

Outre des projections, type COR, dans le cadre de la législation actuelle le modèle Pablo a pu être utilisé pour faire des simulations notamment de mesures *ex ante* de politiques publiques :

- Impact de l'augmentation de la durée de référence par génération ;
- Relèvement des bornes d'âges notamment de la date d'ouverture des droits ;
- Impact de la suppression de chaque bonification ou de l'ensemble des bonifications ;
- Impact de la suppression de la catégorie active.

V) Le système universel de retraite et le développement d'un nouveau modèle

Le SRE a été sollicité dans le cadre des travaux liés à la mise en place d'un système universel de retraites. Le passage à un mode de calcul par points et l'intégration des primes dans l'assiette de cotisations, et donc dans le montant de la pension, étaient les principales nouveautés qui n'étaient pas prises en charge dans Pablo. La question s'est alors posée : faut-il introduire de nouveaux

développements dans Pablo ou construire un nouvel outil ? C'est la seconde voie qui a été choisie tout en cherchant à tirer profit au maximum des potentialités de Pablo. La raison est essentiellement technique : Pablo a été développé en SAS, les possibilités de programmation de ce langage sont limitées et l'exécution est lente. Le nouvel outil, appelé Oscar, est développé en langage R. L'objectif était de pouvoir adapter facilement les programmes aux évolutions fréquentes, ce qui s'est avéré très utile, et d'améliorer nettement les temps d'exécution, ce que l'expérience a confirmé.

En terme de fonctionnalités, un objectif recherché et atteint par Oscar est de pouvoir reprendre aussi bien d'une part des *carrières types* calculées à partir de grilles indiciaires pour des générations, des âges de début de carrières et de liquidation contrôlés par le paramétrage que d'autre part des carrières simulées par le modèle Pablo issues des *bases projetées*. Et comme les fonctions sont unifiées, on assure qu'on obtient les mêmes résultats pour un individu réel de la base projetée qui aurait la même carrière qu'un cas type. Cela permet également de mesurer plus facilement l'impact des différents paramètres de la simulation.

Concrètement, pour les bases projetées, on reprend entre autres la carrière indiciaire, les quotités, les états de chaque individu, le ministère d'appartenance tel qu'ils ont été calculés par Pablo. On impute ensuite un grade qui va fournir un taux de prime. Ce taux de prime, par grade, ne sera pas simulé par le modèle Oscar à chaque période mais maintenu constant tout au long de la carrière de l'agent. Le premier module d'Oscar va alors calculer un salaire mensuel complet, primes incluses : contrairement au modèle carrière de Pablo qui utilise un pas trimestriel, Oscar fonctionne sur un pas mensuel. Ce choix a été fait pour avoir une unité cohérente de temps et ne pas mélanger des dates précises au jour, des âges en mois, des durées ou des états en trimestres. Dans Oscar, chaque événement se passe un mois donné et cela permet une plus grande souplesse pour les prises en compte des grandeurs exogènes comme les évolutions du point FP ou les revalorisations de pensions qui ont toujours lieu en début de mois mais pas toujours de trimestres. Les bonifications ou les naissances des enfants sont également traitées avec ce pas mensuel. Toujours dans un souci de cohérence, Oscar cherche à calculer le maximum de variables par *état* : par exemple les *durées d'assurance autres régimes* sont obtenues par sommation des mois pour lesquels le fonctionnaire se trouve dans un état *autre régime*. C'est un changement par rapport à Pablo où ces durées ont en général été imputées directement. Dans Oscar, on a créé une procédure qui reprend cette durée Pablo (Oscar n'est pas capable de modéliser les carrières ou les états) et la transforme en *état* sans toutefois permettre de remonter avant les 18 ans (16 ans dans le cas des carrières longues) de l'individu. Ce passage en état a été rendu nécessaire car, dans le cadre de la réforme, il fallait pouvoir évaluer des durées en cours de carrière, par exemple à la date prévue pour la transition.

Après avoir calculé un salaire complet, Oscar calcule les cotisations du nouveau système en séparant les montants salarié/employeur mais également contributif/solidarité. Puis la cotisation contributive est transformée en points selon une valeur d'achat du point définie mensuellement.

Oscar a besoin ensuite une date de liquidation : plusieurs options sont alors possibles. Pour les carrières types, on définit une plage d'âges de liquidation, par exemple 62-70 ans, et pour chaque cas type on calcule ainsi les différents montants de pensions aux différents âges. Pour la base projetée, l'option principalement retenue consiste à reprendre la date de liquidation calculée par Pablo. Tout récemment, un modèle de comportement prenant en entrée différents paramètres, dont le montant de la

pension calculé mensuellement sur le même modèle que les carrières types, a été implémenté dans Oscar et permet d'aller plus loin que le modèle de départ au taux plein.

Autre apport d'Oscar : il intègre les cotisations et le calcul de la pension RAFP pour faciliter les comparaisons avec la pension en points qui intègre les primes.

Oscar implémente ensuite les moteurs de liquidation *système actuel* et *système en points*. Concernant le moteur système actuel d'Oscar, l'objectif est d'en faire petit à petit le moteur de référence pour les travaux statistiques du SRE en raison de son aspect modulaire. Il a ainsi pu être testé avec succès sur le flux de liquidation réel 2018. Le moteur de liquidation *système en points* est plus simple : le montant de base est fourni par le produit des points par la valeur de service mensuelle. Toutefois le calcul de la décote/surcote cible, non finalisée dans le projet de réforme à ce jour, reste potentiellement complexe.

Enfin Oscar implémente, comme Pablo, la phase *retraite* : à partir du montant à la liquidation, on applique, jusqu'au décès de l'ayant-droit, les revalorisations, éventuellement différenciées selon le montant. Par contre, contrairement à Pablo, Oscar n'implémente pas la réversion.