Recherches et Prévisions

# MYRIADE : le modèle de microsimulation de la CNAF [Un outil d'évaluation des politiques sociales]

## Un outil d'évaluation des politiques sociales

François Legendre

**Résumé**

Myriade est le modèle de microsimulation en cours de développement à la Caisse nationale des allocations familiales (CNAF). Avec cet outil, la branche Famille améliore ses possibilités d'études, de prévisions et de chiffrages. En effet, les autres sources statistiques dont elle dispose sont représentatives de la population des allocataires et fournissent une information détaillée pour chacun d'eux. Cependant, ces caractéristiques posent parfois problème, notamment lorsqu'il s'agit d'étudier la mise en place d'une nouvelle prestation ou la modification de la législation dans le sens d'un élargissement du champ des bénéficiaires. MYRIADE évite ces difficultés car il est représentatif de l'ensemble des ménages de la France métropolitaine. L'apport de MYRIADE est de trois ordres: un chiffrage peut être fait en prenant en compte l'ensemble des ménages français; chaque individu peut être localisé à la fois dans un ménage, un foyer fiscal ou une famille; MYRIADE permet de distinguer les conséquences fiscales et sociales d'une mesure.

**Citer ce document / Cite this document :**

Legendre François. MYRIADE : le modèle de microsimulation de la CNAF [Un outil d'évaluation des politiques sociales]. In: Recherches et Prévisions, n°66, 2001. Évaluer les politiques familiales et sociales. MYRIADE, un outil de microsimulation à la CNAF. pp. 33-50;

doi : https://doi.org/10.3406/caf.2001.977;

https://www.persee.fr/doc/caf_1149-1590_2001_num_66_1_977;

Fichier pdf généré le 27/02/2024

# MYRIADE : le modèle de microsimulation de la CNAF

## *Un outil d'évaluation des politiques sociales*

François Legendre *, Jean-Paul Lorgnet ** et Florence Thibault ***

***Myriade est le modèle de microsimulation en cours de développement à la Caisse nationale des allocations familiales (CNAF). Avec cet outil, la branche Famille améliore ses possibilités d'études, de prévisions et de chiffrages. En effet, les autres sources statistiques dont elle dispose sont représentatives de la population des allocataires et fournissent une information détaillée pour chacun d'eux. Cependant, ces caractéristiques posent parfois problème, notamment lorsqu'il s'agit d'étudier la mise en place d'une nouvelle prestation ou la modification de la législation dans le sens d'un élargissement du champ des bénéficiaires. MYRIADE évite ces difficultés car il est représentatif de l'ensemble des ménages de la France métropolitaine. L'apport de MYRIADE est de trois ordres : un chiffrage peut être fait en prenant en compte l'ensemble des ménages français ; chaque individu peut être localisé à la fois dans un ménage, un foyer fiscal ou une famille ; MYRIADE permet de distinguer les conséquences fiscales et sociales d'une mesure.***

A partir des bases de données administratives des caisses d'Allocations familiales (CAF), la Caisse nationale des allocations familiales (CNAF) élabore deux produits statistiques qu'elle utilise pour réaliser ses travaux d'études, de prévisions et de chiffrages. Le premier – le FILEAS (fichier des prestations légales et d'action sociale) – fournit une image de la population des allocataires à un instant donné, deux fois par an. Il comprend tous les allocataires et peu d'informations pour chacun d'entre eux. Le second – l'échantillon national des allocataires – est extrait une fois par an. Il contient 1 allocataire sur 52 mais comprend de nombreuses informations pour chacun d'entre eux (Amrouni et Thibault, 2001).

Ces deux sources statistiques sont très précieuses pour la branche Famille de la Sécurité sociale. Cependant, elles ont des limites inhérentes aux bases administratives dont elles découlent. Elles sont notamment tributaires de la législation en vigueur (baisse du nombre de familles durant la période de mise sous conditions de ressources des allocations familiales, etc.) ; elles sont représentatives de l'ensemble des allocataires mais ne le sont pas de l'ensemble des ménages, etc. Par exemple, il n'est pas possible de les utiliser pour chiffrer une augmentation du plafond de ressources d'une prestation car les familles dont les ressources dépassent l'ancien plafond ne sont pas présentes dans les bases administratives.

Pour pallier ces limites, la CNAF a décidé de se doter d'un nouvel outil : un modèle de microsimulation. Ce dernier doit lui permettre, d'une part, de renforcer son expertise pour chiffrer des mesures nouvelles de politiques familiales et, d'autre part, de

* Laboratoire Erudite (Equipe de recherche sur l'utilisation des données individuelles et temporelles en économie - université Paris-XII) et CNAF - Bureau des Prévisions.
** CNAF - Bureau des Prévisions.
*** CNAF - Bureau des Prévisions et Laboratoire CEPN (université Paris-XIII et CNRS).

mieux pouvoir évaluer dans leur ensemble les politiques sociales. MYRIADE est le nom qui a été donné à ce modèle. Il ne s'agit pas d'un acronyme. Ce nom symbolise le très grand nombre d'individus – 220 000 actuellement – dont le modèle cherche à reproduire la situation économique et sociale.

## Un outil général complémentaire aux études

Les méthodes de microsimulation, quand elles sont relatives aux politiques sociales, s'attachent, pour chaque individu d'un échantillon représentatif, à reproduire les principaux événements d'intérêt. Elles permettent, entre autres, de mesurer les conséquences redistributives de ces politiques publiques. De plus, MYRIADE a pour ambition de constituer un outil général, complémentaire aux études, qui aurait pour objet de mieux caractériser une sous-population particulière et/ou un instrument spécifique du système de redistribution.

Pour répondre aux besoins de la CNAF, le modèle est décliné en deux versions. La première version est statique ; elle a trois objets principaux. Tout d'abord, repérer des populations particulières mal appréhendées par d'autres sources statistiques ; ensuite, constituer un outil complémentaire pour prévoir le coût budgétaire des principales prestations des caisses d'Allocations familiales (CAF) ; enfin, chiffrer *ex ante* des mesures nouvelles. La seconde version est dynamique. Elle est utilisée pour évaluer les politiques sociales. La construction du modèle a été entreprise au début de l'année 2000. Fin mars 2001, la version statique du modèle a été livrée. La version dynamique sera développée au cours de l'année 2002. L'objet de cet article est de proposer une première présentation de la version statique du modèle.

Les objectifs que l'on assigne à un modèle de microsimulation conditionnent le choix de son enquête de départ. Par exemple, DESTINIE, le modèle de microsimulation dynamique de l'INSEE, a pour objet la projection à long terme de la situation des retraités. Pour cette raison, ce modèle repose sur l'enquête Actifs financiers de l'INSEE car celle-ci permet de reconstituer la carrière d'activité des individus ; elle est donc particulièrement adaptée. Pour MYRIADE, dédié à l'étude des politiques fiscales et sociales, plusieurs possibilités ont été envisagées.

## L'apport de l'enquête Revenus fiscaux des ménages de l'INSEE

Le tableau 1 recense les enquêtes de l'INSEE qui auraient pu être mobilisées pour servir de base à MYRIADE. Certaines sont utilisées par d'autres modèles de microsimulation existants (1). MYRIADE s'appuie sur l'enquête Revenus fiscaux des ménages de l'INSEE. Cette enquête constitue une source de choix. Elle rapproche les informations de l'enquête Emploi de l'INSEE de données qui proviennent d'une source administrative, les déclarations à l'impôt sur le revenu de la Direction générale des impôts (DGI). L'enquête Revenus fiscaux est annuelle. Aussi, le modèle dispose-t-il d'une population relativement à jour. Cela évite d'avoir à faire « vieillir » l'enquête sur une trop longue période pour cadrer avec la conjoncture de l'année de base.

La taille de l'enquête Revenus fiscaux des ménages – la plus importante réalisée par l'INSEE – permet de limiter les fluctuations d'échantillonnage même pour des populations particulières de taille plus réduite. L'information relative aux dépenses des individus y est toutefois absente. Il n'est donc pas possible d'évaluer la taxe sur la valeur ajoutée (TVA), la taxe intérieure sur les produits pétroliers (TIPP) ainsi que les accises sur le tabac et les alcools acquittés par les ménages.

L'INSEE constitue l'enquête Revenus fiscaux des ménages depuis 1956. Quinquennale jusqu'en 1996, cette enquête est devenue annuelle en raison de l'adoption d'une nouvelle méthodologie. Jusqu'en 1996, un échantillon était tiré par l'INSEE dans le fichier du recensement de la population, échantillon pour lequel

(1) Voir l'article *Les modèles socio-économiques de microsimulation. Panorama et état des lieux pour la France* dans ce numéro de ***Recherches et Prévisions*** (p. 11).

**Tableau 1 - Principales enquêtes sur les personnes en France**

| Enquête | Taille | Périodicité | Principales informations | Institut |
|---|---|---|---|---|
| Patrimoine des ménages | 10 000 ménages | Quinquennale | Description du patrimoine et cycle d'activité<br>Revenus du patrimoine et d'activité | INSEE |
| Budget des familles | 10 000 ménages | Quinquennale | Revenus et dépenses des ménages | INSEE |
| Logement | 30 000 ménages | Tous les 4 ans environ | Description du logement<br>Revenus annuels | INSEE |
| Partie française du panel européen des ménages | 13 000 individus de 17 ans et plus | Les individus sont suivis plusieurs années | Dynamique d'emploi et de ressources des personnes : calendrier d'activité mensuel<br>Ressources annuelles | INSEE |
| Emploi | 75 000 ménages | Annuelle. Les logements sont suivis 3 ans | Activité professionnelle. Calendrier d'activité mensuel. Salaire | INSEE |
| Revenus fiscaux des ménages | 1996 : 25 000 ménages<br>1997 : 50 000 ménages<br>1998 : 75 000 ménages | Annuelle | Par appariement de fichiers administratifs, l'enquête Emploi est « enrichie » d'une information supplémentaire : les revenus déclarés à l'impôt sur le revenu | Direction générale des impôts et INSEE |

Source : François Legendre, Jean-Paul Lorgnet et Florence Thibault, bureau des Prévisions, CNAF, 2001.

la DGI recherchait puis communiquait les revenus fiscaux. Un avis défavorable ayant été donné par la Commission informatique et liberté concernant cette méthode, la recherche des revenus fiscaux est désormais menée directement par l'INSEE qui dispose aujourd'hui des données exhaustives de la DGI. L'INSEE recherche les déclarations fiscales des individus présents dans l'enquête Emploi. Toutefois, certains revenus sont absents : les revenus imposés directement à la source, tels les prélèvements libératoires appliqués aux revenus d'obligations ; les revenus imposables mais situés dans des enveloppes fiscales elles-mêmes exonérées, tels par exemple les revenus des plans d'épargne en action ; enfin, les revenus exonérés comme les revenus des livrets A. De plus, les prestations sociales et les minima sociaux qui ne sont pas imposables à l'impôt sur le revenu ne sont pas connus de l'enquête. Ils sont alors imputés sur barème pour certains, par méthodes économétriques pour d'autres. Enfin, les prélèvements directs sont connus, l'impôt sur le revenu bien sûr mais aussi la taxe d'habitation.

L'enquête Emploi apporte une connaissance détaillée de l'activité des individus. Elle comprend aussi un calendrier d'activité mensuel. De telles informations permettent, par exemple, de calculer le taux de salaire horaire ou de connaître les conditions d'activité des parents de jeunes enfants. Une étude de Jérôme Accardo et de Madior Fall (1996) montre que, de toutes les enquêtes présentées dans le tableau 1, l'enquête Revenus fiscaux des ménages est celle qui couvre le mieux les revenus – du moins les revenus imposables.

## Une enquête qui présente plusieurs avantages

Le choix de l'enquête Revenus fiscaux des ménages pour MYRIADE est dicté par trois raisons principales :

• la première est la qualité des données. Les données de l'enquête Emploi fournissent une description détaillée de l'activité d'un individu sur une année. Les revenus de l'enquête Revenus fiscaux des ménages viennent avantageusement compléter cette source. Connaître l'emploi d'un individu sans apprécier les ressources qu'il en tire ou le contraire serait insuffisant. La complémentarité des deux sources assure au modèle une réelle

capacité à retracer la situation économique et sociale des individus ;

• la deuxième raison est la taille de l'échantillon. La richesse qualitative de l'enquête est pleinement mise à profit du fait d'une richesse quantitative : l'échantillon 1998 est déjà constitué de 220 000 individus et le prochain en comptera 300 000. L'étude de certaines populations nécessite un tel effectif. La robustesse et la significativité des résultats en dépendent ;

• enfin, la dernière raison est la fréquence de l'enquête. Parmi les enquêtes disponibles, elle est la seule à avoir un pas annuel. Le modèle est ainsi doté d'une population dont les caractéristiques sont régulièrement renouvelées.

## Le système d'information du modèle MYRIADE

Le système d'information de MYRIADE est caractérisé par trois points saillants (2) :

• *l'unité élémentaire est l'individu*. Chaque individu est repéré par un identifiant unique, qui permet l'accès à ses caractéristiques mais aussi à sa désignation dans la représentation des liens entre les individus. Cette représentation des liens est très peu contraignante : le nombre d'enfants n'est pas limité, les familles recomposées peuvent être représentées sans restriction, etc. ;

• *quatre regroupements sont distingués : le ménage, le foyer fiscal, la famille au sens des prestations familiales et la famille au sens des minima sociaux*. Ces regroupements constituent quatre partitions complètes de la population, qui viennent s'ajouter à la partition triviale des individus. Le ménage est une entité utilisée à l'INSEE. Il est défini comme l'ensemble des personnes vivant dans un même logement. A chaque ménage est associée une personne de référence unique. Le foyer fiscal est une entité propre à la DGI. C'est l'ensemble des personnes qui figurent sur une même déclaration d'impôt sur le revenu. Pour chaque foyer fiscal, il existe un « déclarant » unique. Les familles, au sens des prestations familiales, sont constituées des parents et des enfants comptés à charge pour de telles prestations. Ce sont les enfants de moins de vingt ans dont les salaires ne sont pas supérieurs à 55 % du SMIC brut. Enfin, la famille au sens des minima sociaux est plus large que la précédente car l'âge limite pour être enfant à charge est porté à 25 ans. Ces quatre regroupements peuvent faire l'objet d'aménagements au gré de l'évolution de la législation ou pour réaliser certaines commandes ;

• *les paramètres qui décrivent la législation et la réglementation sont codés dans un fichier externe de configuration*. Concrètement, ce fichier est un tableau au format texte. Une ligne du tableau commence par le nom d'un paramètre [montant du SMIC brut horaire, plafond de base de l'allocation pour jeune enfant (APJE)]. Suivent ensuite ses valeurs pour les différentes années. La mise à jour de ces tableaux est relativement aisée à l'aide d'un tableur. Une seule contrainte existe : la sauvegarde doit impérativement être réalisée au format texte. Lorsque le programme de microsimulation est activé, les paramètres sont lus et mémorisés dans des variables internes au modèle. Par exemple, dans les programmes de recalcul des prélèvements obligatoires et des transferts sociaux, les valeurs prises par ces paramètres ne sont pas « codées en dur » ; ils prennent la valeur présente dans le fichier texte disponible : il est donc très facile de chiffrer les mesures nouvelles en modifiant les valeurs des paramètres dans ce fichier de configuration.

## L'année est la période de référence du modèle

Dans MYRIADE, le pas temporel de simulation – la période de référence – est l'année, pour une raison essentielle : les données qui proviennent de la source fiscale sont des grandeurs de flux, relatives à l'année précédente, et on ne veut ni se

(2) On a été contraint de développer le propre système d'information de MYRIADE, non par désir de faire compliqué mais parce que l'on a très tôt mesuré les limites des logiciels habituellement utilisés par les économistes comme, par exemple, le système SAS [voir (4) p. 39]. On a alors décidé de développer MYRIADE en utilisant un langage de programmation, le langage C++.

**La réalisation d'une microsimulation**

Le processus de la réalisation effective d'une microsimulation, pour l'année *N*, est le suivant :

*1. Lecture dans sa totalité du fichier de la population depuis le support externe* – le disque dur du micro-ordinateur. La mémoire centrale de l'ordinateur contient la représentation de toute la population pour l'année *N – 1*. Cette mémoire doit donc être de taille suffisante – au moins 128 méga octets.

*2. Réalisation effective de la microsimulation pour l'année N*. Cette étape est très rapide. Les informations individuelles nécessaires sont toutes présentes dans la mémoire centrale. Elles sont mises à jour « sur place » dans cette mémoire ; les nouvelles données de l'année *N* engendrées par la simulation prennent la place des anciennes données de l'année *N – 1*.

*3. Ecriture de la totalité de la population dans un nouveau fichier* – sur le support externe. Ce fichier pourra être relu plus tard par le système SAS (1), au moyen d'une interface, pour produire les chiffres d'intérêt (états détaillés, statistiques par décile, etc.).

*4. Soit retour à la deuxième étape soit arrêt du traitement*. Le retour à la deuxième étape permet d'enchaîner la simulation sur plusieurs années, pour l'année *N+1, N+2*…

Cette conception permet de réduire les temps de réponse et ne peut être développée qu'au moyen d'un langage de programmation. Les temps de réponse du modèle sont faibles, car les données résident en mémoire centrale. Pour donner un ordre d'idées, une simulation pour une année est réalisée en une dizaine de secondes. Pendant ce laps de temps, le modèle MYRIADE calcule et sauvegarde l'impôt sur le revenu de tous les foyers fiscaux, la taxe d'habitation de tous les ménages, les prestations familiales et les minima sociaux de toutes les familles, les cotisations patronales et salariales de l'ensemble de la population. Dès lors, il est possible de réaliser des simulations pour plusieurs années ou d'exécuter des variantes sans que les temps de réponse ne deviennent rédhibitoires.

Ce système d'information présente toutefois un triple inconvénient. Il exige des compétences en C++. Le coût d'entrée est élevé et limite en cela l'accès au modèle pour un utilisateur occasionnel. Enfin, il rend le modèle plus difficile à transmettre. Toutefois, MYRIADE est convenablement documenté.

(1) SAS est un progiciel qui facilite la gestion de données d'enquête et leur exploitation. Il est beaucoup utilisé par les statisticiens pour dépouiller des enquêtes, obtenir des états, réaliser des traitements statistiques, etc.

priver de cette information *a priori* robuste ni la modifier. Par ailleurs, pour la plupart des prestations familiales sous condition de ressources, les montants de la déclaration fiscale correspondent bien aux variables de la déclaration de ressources que les allocataires déposent dans leur CAF. On est ainsi en mesure de simuler très précisément ces prestations familiales. Pour autant, la périodicité pertinente pour certains éléments du système fiscalo-social n'est pas l'année : les cotisations sociales sont assises sur des flux mensuels, la condition de ressources de la plupart des minima sociaux est basée sur un calendrier trimestriel, et leur versement est mensuel, etc.

Par conséquent, on pourrait souhaiter décomposer chaque poste de la déclaration fiscale en données mensuelles et se contenter, en première approximation, de diviser par douze son salaire annuel. Dans le cas d'un individu qui est salarié six mois, chômeur cinq mois et en arrêt maladie le dernier mois, le calcul est beaucoup plus complexe. Ses salaires, ses indemnités de chômage, ses indemnités journalières de maladie et ses primes éventuelles sont agrégées dans un même revenu catégoriel de la déclaration à l'impôt sur le revenu.

La question est alors de savoir comment éclater ce poste entre ces différentes composantes. D'abord, les données relatives à l'enquête Emploi sont pour la plupart annuelles, à l'exception du calendrier d'activité qui est mensuel. Le passage à un pas mensuel ou trimestriel ne pourrait donc se faire qu'en retenant des hypothèses fortes et le gain en information en serait

limité. Ensuite, le calendrier d'activité de l'enquête Emploi informe sur le nombre de mois salariés et le nombre de mois chômés de cet individu mais pas sur les montants perçus chaque mois.

L'enquête Emploi est simplement renseignée sur la perception d'allocations chômage : on connaît le montant de la dernière allocation versée ainsi que le nombre de jours couverts par celle-ci. Ces données ne permettent pas de reconstituer les gains mensuels de l'individu. Il faudrait faire des hypothèses sur le nombre de mois pendant lesquels l'allocation chômage a été versée, sur la dégressivité de celle-ci, sur le montant des primes et des indemnités journalières, etc. L'exercice est plus difficile si l'individu est chômeur, puis salarié lors de l'interview à l'enquête Emploi. En effet, les allocations chômage éventuelles ne seront pas renseignées dans ce cas. Cet exemple éclaire sur les difficultés rencontrées pour créer un pas infra-annuel (3).

## MYRIADE retient un seul statut d'occupation dans l'année

Pour que MYRIADE puisse rester (relativement) simple, il fallait fortement réduire cette complexité. L'option retenue est double. D'une part, on affecte à chaque individu du modèle un et un seul statut d'occupation dans l'année et, d'autre part, on simule dans le modèle les prélèvements et les transferts pour un laps de temps annuel. On ne s'est pas engagé sur la voie d'envisager par individu plusieurs statuts d'occupation dans l'année ni sur celle de retenir un pas infra-annuel de simulation.

Le calendrier rétrospectif du statut d'occupation mensuel de l'enquête Emploi permet de déterminer le « statut d'occupation dominant » dans l'année. Pour MYRIADE, un individu, par exemple, est considéré comme chômeur si ce statut est celui qui présente la plus forte occurrence dans l'année. Deux approximations sont ainsi émises :

• en premier lieu, le revenu annuel porté dans la déclaration fiscale est apprécié comme relevant en totalité du statut d'occupation dominant. Par exemple, le revenu d'un individu, dont le statut dans MYRIADE est salarié, est traité dans son intégralité comme un salaire. A ce titre, il supporte les cotisations sociales, la contribution sociale généralisée (CSG), etc. ;

• en second lieu, la législation est, quand cela est nécessaire, « annualisée ». Par exemple, la condition de ressources du revenu minimum d'insertion (RMI) porte sur un revenu trimestriel ; dans MYRIADE, elle porte sur le montant annuel correspondant.

Les conséquences ne sont pas neutres. En retenant ce procédé, on affecte à tous les bénéficiaires du RMI un revenu disponible annuel égal au plafond de ce minimum social, augmenté du montant des allocations logement. Par ailleurs, le RMI ne sera pas simulé pour un individu qui aurait des ressources annuelles supérieures au même plafond (même s'il pourrait en bénéficier sur une partie de l'année). Enfin, les mécanismes d'intéressement ne sont pas simulés. On contrôle, du mieux possible, les conséquences de ce parti pris. Quand on fait « vieillir » les données, on se cale sur des grandeurs obtenues grâce à l'enquête Emploi de l'année de base du modèle en recourant à ce même calcul de « statut d'occupation dominant ». De même, pour évaluer le nombre de bénéficiaires du RMI avec la condition de ressources « annualisée », on décompte, dans l'échantillon national des allocataires de la CNAF, le nombre d'individus qui bénéficient, dans ce schéma, de ce minimum social.

## MYRIADE retraite l'enquête Revenus fiscaux

Le retraitement de l'enquête Revenus fiscaux des ménages de l'INSEE a un double objectif : d'une part, rendre les données propres à constituer l'échan-

(3) D'autres modèles de microsimulation ont cependant fait le choix de mensualiser les revenus.

tillon de base du modèle et, d'autre part, faire « vieillir » l'enquête afin qu'elle retrace le contexte de l'année de base du modèle. Actuellement, on utilise la dernière enquête Revenus fiscaux disponible, celle de l'enquête Emploi de mars 1998 couplée aux déclarations à l'impôt afférentes aux revenus de 1997. En revanche, l'année de base du modèle est l'année 1999 ; ainsi, la microsimulation peut être enchaînée pour 2000, 2001, etc. Ce retraitement comporte deux étapes.

### *La première étape comporte un quadruple objectif*

La première étape, réalisée en SAS (4), a un quadruple objet.

· ***La reconstitution des foyers fiscaux au sein des ménages***. L'enquête Revenus fiscaux est destinée à produire des connaissances sur les revenus des ménages et des individus. La plupart des études menées à partir de cette enquête produisent des résultats dont l'unité est soit le ménage soit l'individu. L'un des objectifs de MYRIADE est de pouvoir mesurer les conséquences redistributives des politiques publiques. Pour mener à bien un tel objectif, le modèle doit pouvoir rendre compte des prélèvements qui pèsent sur les individus, notamment l'impôt sur le revenu. Ceci n'est possible qu'avec la création d'une entité propre à la DGI : le foyer fiscal. Un ménage peut être composé de plusieurs foyers fiscaux. C'est le cas notamment des ménages de concubins, puisque chaque concubin constitue un foyer fiscal ; s'ils ont des enfants il faut les rattacher au foyer fiscal de l'un des parents.

· ***La préparation des variables de MYRIADE***. On construit les variables du modèle, à partir soit des informations de l'enquête Emploi, soit des données de la déclaration fiscale. Par exemple, on distingue, pour les individus en âge de travailler, six statuts d'activité : inactif, chômeur, salarié « précaire », salarié du privé, salarié du public et indépendant. Ceci est un choix, il existe beaucoup plus de statuts d'activité dans l'enquête Emploi. A chaque fois, il faut arbitrer entre le « bon » degré de détails et le « bon » degré de simplification propre à toute modélisation.

Plus généralement, la réalisation d'un tel modèle nécessite de la part des concepteurs une recherche constante de l'efficacité. Celle ci se traduit par des arbitrages dont on essaie toujours de rendre compte. Les variables relatives à la taxe d'habitation sont construites à ce stade du développement du modèle. Pour chaque ménage, il s'agit de déterminer les taux d'imposition communaux, départementaux et régionaux et les bases des abattements de taxe d'habitation.

· ***Le calage sur la conjoncture de l'année de base***. On redresse la pondération des ménages de l'enquête pour prendre en compte le contexte macro-économique de l'année de base. A côté de toutes les variables qui servent à caler l'enquête Revenus fiscaux, le premier critère est constitué par les statuts d'activité en trois modalités (inactif, chômeur, occupé) croisés avec le sexe et avec trois tranches d'âge ; le deuxième critère de calage est la taille des ménages en quatre modalités. Viennent ensuite des critères telles la région de résidence et la catégorie socioprofessionnelle de la personne de référence du ménage. A ce stade, les grandes masses de revenus, traitements et salaires, pensions et retraites et les revenus d'indépendants sont calés sur les chiffres de la DGI.

· ***La préparation d'un fichier de transfert***. Il faut mettre les données au format du modèle. Pour éviter toute erreur de transfert, une interface automatique, écrite en langage SAS, génère un fichier de transfert et le fragment de programme C++ qui peut relire ce fichier de transfert. Lors de la lecture de ce fichier, de nombreux tests sont exécutés, qui valident la préparation. Lors de ces tests, certaines erreurs sont bloquantes, d'autres ne le sont pas, mais dans ce cas

(4) SAS est un progiciel qui facilite la gestion de données d'enquête et leur exploitation. Il est beaucoup utilisé par les statisticiens pour dépouiller des enquêtes, obtenir des états, réaliser des traitements statistiques, etc.

l'utilisateur est averti dans le journal du transfert.

### *La seconde étape poursuit six objectifs*

La seconde étape de ce retraitement, réalisée en C++, poursuit six objectifs.

· ***Le clonage des ménages***. MYRIADE repose sur une représentation à taux constant de la population française de Métropole ; il est attribué à chaque individu un même facteur de pondération. En revanche, dans l'enquête Revenus fiscaux, les individus ont des poids différents. Autrement dit, l'INSEE attribue à chaque ménage un poids en fonction du nombre de ménages ayant les mêmes caractéristiques, qui sont présents dans la population totale. Une procédure de clonage, c'est-à-dire de répétition à l'identique mais en nombre variable des ménages, permet de passer de l'une à l'autre de ces deux représentations.

Pour l'enquête de 1997, le nombre initial d'individus est de 111 814 ; il est de 221 125 dans MYRIADE, le facteur moyen de clonage étant de l'ordre de deux. Ce facteur moyen de clonage suffit pour assurer une représentativité satisfaisante de l'échantillon à taux constant. C'est pourquoi dans MYRIADE chaque individu a le même poids. La future approche dynamique du modèle imposait un taux constant. Par exemple, si on veut modéliser des mises en couples dans le modèle, il est nécessaire de pouvoir assembler des individus de même poids. On ne peut envisager de marier un individu dont le poids est cent avec un autre dont le poids est trois cents.

· ***Le rapprochement des foyers fiscaux au sein de ménages distincts***. Cette étape complète la reconstitution des foyers fiscaux au sein des ménages. On rapproche, statistiquement, des membres de ménages distincts qui appartiennent en fait au même foyer fiscal. En effet, on peut repérer des foyers fiscaux « incomplets » que l'on rapproche alors de ménages pour lesquels aucune déclaration fiscale n'a été retrouvée. Les enfants ayant décohabité, mais qui sont rattachés fiscalement au foyer de leurs parents, ont été réintégrés dans le foyer fiscal parental. A l'issue de ce traitement, chaque individu est présent dans un et un seul foyer fiscal, de sorte que l'ensemble des foyers fiscaux réalise une partition totale de l'ensemble des individus.

· ***Le calage sur la pyramide des âges par sexe de l'année de base***. Pour correctement simuler un grand nombre de prestations familiales, il faut être très précisément calé sur la pyramide des âges de l'année de base. Ceci est obtenu en altérant les âges de certains individus de la population sans toutefois modifier les dates de naissance de plus ou de moins d'une année.

· ***L'évolution des revenus depuis l'année de l'enquête***. On fait « vieillir » les revenus de l'enquête en appliquant les évolutions de la comptabilité nationale.

· ***L'imputation sur barème des revenus bruts***. L'enquête Revenus fiscaux concerne les revenus imposables, nets des cotisations sociales, exceptée la contribution sociale généralisée non déductible. Par exemple, dans le cas des salaires, il faut « remonter » depuis le salaire net au coût du travail. La détermination du revenu super brut (composé des revenus d'activités ainsi que des cotisations salariales et patronales) à partir du revenu déclaré au fisc est faite sur barème, à l'aide des statuts d'occupation précédemment incorporés dans le modèle.

· ***L'étalonnage d'un certain nombre de transferts sociaux***. L'éligibilité à certains transferts sociaux n'est qu'imparfaitement observée dans les données. Ainsi, l'allocation parentale d'éducation (APE) et l'allocation aux adultes handicapés (AAH) font l'objet d'un tel retraitement où la condition d'éligibilité est redressée afin que le nombre total de bénéficiaires de ces deux allocations soit exactement calé. D'autres prestations, tels le minimum vieillesse et le RMI, font l'objet d'un

non-recours significatif ; le modèle, spontanément, simulerait « trop » de bénéficiaires pour ces prestations.

Le non-recours au RMI peut être dû au caractère stigmatisant de la prestation. Le minimum vieillesse, servi à une personne dont la succession future se révélerait être supérieure à 38 112 euros, sera imputé sur le montant de la succession. Une partie du non-recours au minimum vieillesse est expliquée par cette clause. Dans tous les cas, cet étalonnage a été réalisé en reproduisant la structure de la population des bénéficiaires, obtenue soit à partir de l'échantillon des allocataires de la CNAF soit, pour le minimum vieillesse, à partir de données de la Direction de la recherche, des études, de l'évaluation et des statistiques (DREES) du ministère de l'Emploi et de la Solidarité.

## Une structure modulaire

MYRIADE est organisé de façon modulaire. Cette organisation permet de développer et de tester des modules indépendamment les uns des autres. Elle permet aussi de décentraliser la construction du modèle et d'identifier précisément les variables que le module utilise en lecture seule et celles qu'il permettra de renseigner. Aussi est-il nécessaire de bien spécifier l'ordre d'activation des modules. La législation détermine partiellement cet ordre. Par exemple, les minima sociaux sont calculés après les prestations familiales dont ils dépendent, la taxe d'habitation est calculée après les minima sociaux afin de tenir compte des diverses exonérations prévues par la législation. Cet ordre détermine la séquence des événements dans l'année : par exemple, pour un actif occupé, tout d'abord la perte de son emploi puis ensuite le retour éventuel à l'emploi.

Fin 2001, six modules sont achevés. Ils retracent le passage du revenu super brut des individus – avant toute cotisation – au revenu disponible des familles – après prélèvements et transferts. Les modules relatifs aux comportements (transition sur le marché du travail, équations de salaire, démographie, mise en couple…) seront introduits dans le courant de l'année 2002.

## La détermination de l'impôt sur le revenu : le module IRPP

Le premier module de MYRIADE est consacré au calcul de l'impôt sur le revenu des personnes physiques. Ce calcul a été légèrement simplifié afin d'alléger la programmation et la maintenance du module. Certaines réductions ou déductions ne sont pas prises en compte et certains dispositifs fiscaux, telle l'imposition suivant le système du quotient ou le système du taux effectif, n'ont pas été traités. Pour le reste, la programmation du module a été réalisée en suivant de près la fiche de calcul proposée par la DGI.

Pour chaque foyer fiscal, chaque type de revenu catégoriel est traité séparément et se voit appliquer les abattements qui lui sont propres. On retranche du revenu brut global l'ensemble des charges déductibles retenues pour obtenir le revenu net global, duquel sont déduits les abattements spéciaux (pour personnes âgées ou enfants à charge ayant fondé un foyer). Le nombre de parts fiscales du foyer est ensuite déterminé. Pour les revenus de l'année 2000, la prime pour l'emploi est calculée. Il est tenu compte du plafonnement du quotient familial, de la décote et des diverses réductions d'impôt.

## Le calcul de la taxe d'habitation du ménage : le module TH

L'enquête sur les revenus fiscaux des ménages comprend, notamment pour chaque ménage, le montant de la taxe d'habitation (TH) net des abattements, dégrèvements et exonérations prévus par la législation, la valeur locative ca-

dastrale brute de l'habitation principale ainsi que la base nette d'imposition à la TH. Les abattements sont pratiqués sur la valeur locative brute. Les dégrèvements et exonérations sont appliqués au montant brut de la taxe d'habitation. Ces diverses mesures profitent aux ménages les moins favorisés ; elles ne sont pas neutres vis-à-vis de la redistribution. Il convient donc de pouvoir les recalculer dans le modèle. C'est l'objet du module TH.

### *Une difficulté : l'absence de données infra-départementales*

Le travail décrit ci-après est réalisé lors de la préparation des données pour MYRIADE. Afin de préserver une unité à la description du calcul de la TH, on le détaille ici. Pour évaluer le montant des abattements, il faudrait pouvoir localiser chaque ménage dans sa commune. Lors de l'élaboration du module TH, la principale difficulté rencontrée est liée à l'absence de données infradépartementales dans l'enquête sur les revenus fiscaux. Ces informations ne sont pas transmises pour rendre les données complètement anonymes. Or la fiscalité locale varie considérablement selon la commune. Pour lever cette difficulté, un taux d'imposition apparent est calculé pour chaque ménage. Au sein du département, ce taux sert à regrouper les ménages par commune sans que cette dernière ne soit identifiée. Par la suite, on calcule en moyenne le montant des abattements propres à chaque collectivité locale : la région, le département, connus de l'enquête, et la commune repérée à l'aide de ce taux d'imposition apparent. Cette méthode préserve l'anonymat, puisque les communes ne sont absolument pas identifiées. De plus cette méthode est robuste, car la structure aréolaire (5) de l'enquête Emploi garantit un nombre limité de communes par département.

Les dégrèvements et les exonérations sont calculés dans le module TH. Ces allègements de charge bénéficient aux personnes âgées, aux bénéficiaires de minima sociaux et aux ménages percevant des revenus modestes. Le module TH doit donc être exécuté après le module relatif aux minima sociaux.

## Le module cotisations sociales

Le module cotisations sociales calcule, pour chaque individu, le montant total des cotisations sociales employeurs et employés ainsi que la contribution sociale généralisée (CSG) et la contribution au remboursement de la dette sociale (CRDS) versées. Les cotisations des risques vieillesse et chômage sont distinguées notamment parce qu'elles ouvrent droit à un revenu individualisable. MYRIADE distingue les agents de l'Etat, les salariés non cadres du privé, les cadres du privé, les chômeurs, les retraités, les agriculteurs et, enfin, les autres indépendants. Pour cette dernière catégorie, les différents régimes des professions libérales ne sont pas isolés. Les allègements de charge sur les bas salaires ont été modélisés. La CSG et la CRDS sur les revenus du patrimoine sont calculées. Les revenus concernés sont les bénéfices fonciers, les revenus de capitaux mobiliers et les plus-values foncières et mobilières. Ce module est complété par le calcul de la CRDS sur les prestations familiales. Enfin, il intègre les exonérations et les réductions de CSG pour les chômeurs et les retraités ayant de faibles revenus.

## De la notion de « ménages » à celle de « familles »

L'enquête Revenus fiscaux distingue trois niveaux de regroupement : l'individu, le foyer fiscal et le ménage. Le regroupement « famille » sur lequel repose l'attribution d'un grand nombre de transferts sociaux ne correspond à aucun d'entre eux. Un ménage peut comprendre plusieurs familles tout comme un foyer fiscal. Une famille peut être composée de plusieurs ménages ou de plusieurs foyers fiscaux. Il est donc indispensable dans MYRIADE de

(5) On entend par structure aréolaire le fait que, sur une même année, toutes les communes ne sont pas enquêtées. Seules quelques aires géographiques sélectionnées de façon aléatoire sur l'ensemble du territoire métropolitain sont utilisées comme base de sondage.

reconstituer les familles avant de procéder au calcul des prestations.

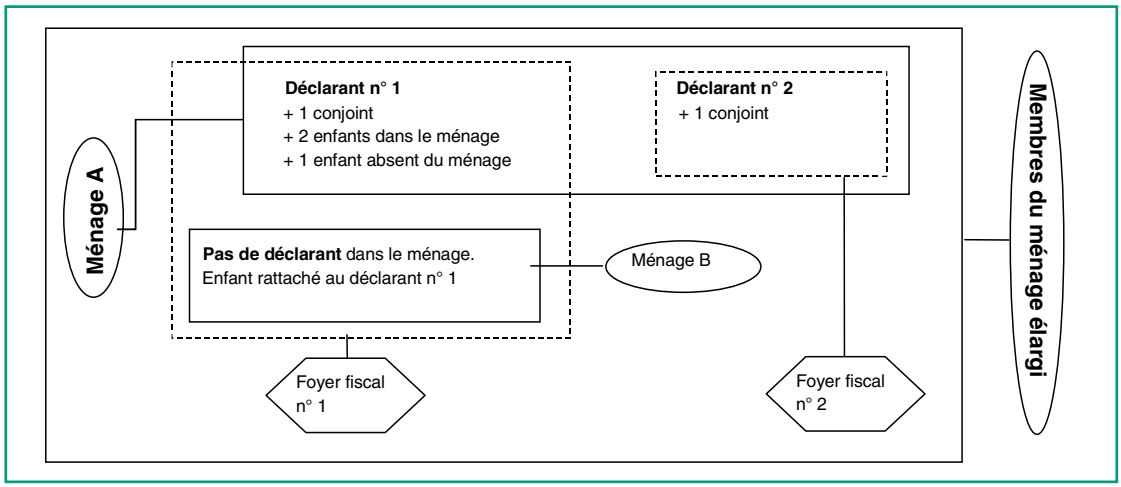
Par ailleurs, la « famille » n'est pas définie de manière unique. Selon la prestation, les personnes susceptibles d'être comptées à charge ou les ressources prises en compte varient. On distingue deux grandes notions de famille qui ont donné lieu à autant de modules : prestations familiales et minima sociaux. L'algorithme retenu pour reconstituer les familles repose sur la constitution d'un « ménage élargi ». On repère tout d'abord tous les déclarants de chaque ménage. Le ménage élargi est l'union des membres de leur foyer fiscal (schéma ci-dessous). Le passage par le foyer fiscal permet notamment de réintégrer les étudiants absents du ménage mais susceptibles d'être comptés comme enfants à charge. Ces ménages élargis constituent une partition complète de la population. Tous les individus du modèle appartiennent à un et à un seul ménage élargi. Enfin, chaque ménage élargi est décomposé en « famille » selon les règles d'éligibilité propres à chaque prestation.

L'ordre dans lequel sont simulées les prestations est dicté par deux règles. La première est la règle du non-cumul de certaines allocations. Pour déterminer celles que la famille « préfère », on a supposé qu'elle faisait d'abord des choix comportementaux, notamment vis-à-vis du marché du travail (l'allocation parentale d'éducation est la première prestation analysée dans le module prestations familiales). La seconde règle déterminant l'ordre de traitement des prestations est la conditionnalité du versement de certaines prestations à la perception de certaines autres. Par exemple, le complément d'allocation aux adultes handicapés suppose la perception d'une allocation logement ou d'une aide personnalisée au logement, de sorte que le module prestation logement est exécuté avant le module minima sociaux.

## Le module des prestations familiales

Comme on l'a mentionné *supra*, la première allocation calculée dans le module des prestations familiales est l'allocation parentale d'éducation. Les prestations sous

**Constitution des membres du ménage élargi**




Source : François Legendre, Jean-Paul Lorgnet et Florence Thibault, bureau des Prévisions, CNAF, 2001.
*Lecture du shéma* : MYRIADE repère dans chaque ménage les déclarants à l'impôt sur le revenu et fait l'union des membres leurs foyers fiscaux. On part du ménage A. Ce dernier comprend deux déclarants : il faut donc repérer leurs membres, même ceux qui sont absents du ménage A. Le foyer fiscal du déclarant n° 1 est composé de lui-même, de son conjoint, de deux enfants présents dans le ménage A et d'un autre enfant – un étudiant, par exemple – qui dispose de son propre logement. Ce dernier constitue à lui seul le ménage B. Pour construire le ménage élargi, il est rattaché au ménage A. Le foyer fiscal n° 2 (les parents du déclarant n° 1, par exemple) est formé du déclarant n° 2 et de son conjoint. Le ménage élargi constitué par MYRIADE comprend au total sept personnes : les deux déclarants, leurs conjoints et les trois enfants du déclarant n° 1.

condition de ressources – qui reposent sur le nombre d'enfants à charge et le nombre d'enfants ouvrant droit à une prestation en raison de leur âge – donnent des résultats très satisfaisants pour l'année de base du modèle. Il s'agit des allocations familiales, de l'allocation pour jeune enfant courte, de l'allocation pour jeune enfant longue, du complément familial et, dans une moindre mesure, de l'allocation de rentrée scolaire.

En revanche, l'allocation de soutien familial n'est qu'imparfaitement appréhendée par MYRIADE car l'enquête Revenus fiscaux permet uniquement de repérer les orphelins donnant droit à une demi-part fiscale supplémentaire, c'est-à-dire ceux issus du mariage et dont le parent veuf n'est pas remarié. D'après l'échantillon national des allocataires de la CNAF de 1998, cette contrainte limite à 26 % le nombre de familles bénéficiaires repérables avec MYRIADE. L'allocation d'éducation spéciale souffre de la même limite. Dans l'enquête Revenus fiscaux, on repère seulement les enfants ouvrant droit à une demi-part fiscale supplémentaire en raison de leur handicap de plus de 80 %, ce qui réduit pratiquement de moitié nombre de familles bénéficiaires de cette prestation dans le modèle.

Les bénéficiaires de prestations subventionnant la garde des jeunes enfants ne sont pas toujours bien identifiés. Pour l'allocation de garde d'enfant à domicile, la présence d'un salarié à domicile est mentionnée dans les déclarations fiscales des familles. En revanche, il n'est pas possible de distinguer les familles recourant à la crèche de celles bénéficiant d'une aide à la famille pour l'emploi d'une assistante maternelle agréée car le montant de ces différents frais de garde est reporté sous la même rubrique de la feuille d'impôt.

## Le module des prestations logement

Le rôle des allocations logement dans le système redistributif national confère au module des prestations logement une importance particulière. Il convenait de rendre compte au plus près de ces allocations. L'ensemble du barème a été reconduit dans MYRIADE et le calcul suit minutieusement la législation. Néanmoins un problème subsiste. MYRIADE ne traite pas de la même façon les aides aux locataires et celles liées à l'accession à la propriété. Dans l'enquête Revenus fiscaux, il n'existe pas d'information sur le montant du loyer ou de la mensualité de remboursement de prêt pour les ménages accédant à la propriété. En revanche, certains renseignements sur le logement sont disponibles, notamment le statut d'occupation (accédant, locataire en HLM, logé à titre gratuit…), le nombre de pièces, le type d'immeuble… Toutefois, les informations sont limitées.

En conséquence, pour les locataires, on a imputé un loyer à chaque ménage en fonction de ses caractéristiques et des informations fournies par l'enquête Revenus fiscaux à partir d'un ajustement économétrique conduit sur l'enquête Logement, et on a (re)calculé le montant de l'allocation sur barème. Toute modification de la situation du ménage sera donc prise en compte par MYRIADE. En revanche, pour les aides à l'accession, les données de l'enquête Revenus fiscaux sont insuffisantes pour imputer à chaque ménage l'ensemble des paramètres nécessaires au calcul de ses droits, notamment les dates et les montants de prêt. Le montant de l'aide sera imputé directement à chaque ménage en fonction de ses caractéristiques. Aussi les prestations logement à l'accession sont-elles figées dans le modèle. Une modification de la législation n'aura donc aucune incidence pour les familles bénéficiaires d'une telle prestation logement.

## Le module des minima sociaux

MYRIADE modélise quatre minima sociaux : l'allocation de parent isolé, le RMI, l'allocation aux adultes handicapés et le minimum vieillesse. Soumises à l'impôt, l'allocation de solidarité spécifique et l'allocation d'insertion sont déjà prises en compte. Pour le RMI et l'allocation de

parent isolé, le pas annuel – à savoir la période de référence sur laquelle sont effectuées la collecte de données du modèle – conduit à utiliser comme référence le nombre de familles ayant perçu la prestation au moins un mois dans l'année et maintenues dans leur droit lorsque ce dernier est calculé à partir de leur déclaration de ressources annuelles, au lieu des déclarations trimestrielles. Leur nombre est recalculé à partir des informations disponibles dans l'échantillon national des allocataires.

Pour ces deux prestations, les effectifs de familles bénéficiaires sont nettement plus élevés sur l'année qu'à l'instant donné (en règle générale les statistiques portant sur les bénéficiaires se réfèrent au 31 juin ou au 31 décembre). Par rapport aux effectifs du 31 décembre 1998, le nombre de bénéficiaires du RMI est plus élevé de 41 % et le nombre de bénéficiaires de l'allocation de parent isolé est plus élevé de 75 %. Cet écart est justifié par les taux d'entrée et de sortie importants dans ces prestations. En revanche, la référence aux ressources annuelles conduit à rejeter 20 % de ces dossiers dans les deux cas.

L'allocation de parent isolé est le premier minima modélisé. On suppose donc que tous les bénéficiaires potentiels sont des bénéficiaires effectifs. On ne modélise cependant pas l'allocation de parent isolé dite « courte » parce que le fait générateur – décès du conjoint ou séparation quand le benjamin a plus de 3 ans – est difficile à repérer dans les données. De même, les femmes enceintes n'ayant pas d'autre enfant à charge ou des enfants à charge de plus de 3 ans ne sont pas directement identifiables. Au total, MYRIADE ne parvient pas à retrouver environ 30 % des bénéficiaires de l'allocation de parent isolé.

L'allocation aux adultes handicapés est ensuite simulée. Selon les ressources et la configuration familiale, elle peut être versée à taux plein ou à taux partiel et concerner la personne de référence de la famille et / ou son conjoint. Pour le minimum vieillesse, on retient les familles dans lesquelles une personne au moins est âgée de plus de 60 ans et a un handicap, ou de plus de 65 ans à l'exclusion des allocataires de l'allocation aux adultes handicapés. La majoration de 10 % accordée pour les personnes ayant eu trois enfants au moins n'est pas prise en compte dans MYRIADE, puisque cette information n'est pas connue.

Le RMI est modélisé en dernier. Pour déterminer le nombre d'enfants à charge, on conserve les enfants de moins de 25 ans ayant des ressources inférieures au montant de la majoration du RMI à laquelle ils peuvent donner droit. Sont ensuite calculées les ressources de la famille en prenant en compte le forfait logement. En revanche, en raison du pas annuel du modèle, il n'a pas été possible d'intégrer les abattements et les neutralisations dont a pu bénéficier la famille à un moment donné. De même, l'intéressement attribué en cas de reprise d'activité n'est pas inclus dans le module.

## La validation des transferts sociaux

Les statistiques de la CNAF (tableau 2, colonnes 1 et 2 p. 46) n'ont pas pu être systématiquement reprises pour servir de référence dans MYRIADE. Trois raisons principales expliquent cette situation. D'une part, les statistiques nationales fournissent des nombres de « familles » bénéficiaires tandis que MYRIADE calcule – lorsque les droits sont liés à des caractéristiques individuelles d'un ou des deux parents – des nombres d'« individus » bénéficiaires. Cette différence joue dans le cas de l'allocation aux adultes handicapés. Une famille bénéficiaire peut comporter deux individus ayants droit si les deux parents répondent aux conditions d'éligibilité de la prestation. D'après l'échantillon national des allocataires de 1998, cette situation concerne 1,2 % des familles bénéficiaires.

**Tableau 2 - Une première validation de MYRIADE pour l'année 2000 en France métropolitaine**

| Intitulé de la prestation et de son unité de référence | Statistiques nationales | | Cibles pour MYRIADE | | Résultats de MYRIADE | |
|---|---|---|---|---|---|---|
| | en millions d'euros | en milliers d'unités | en millions d'euros | en milliers d'unités | en millions d'euros | en milliers d'unités |
| Bénéficiaires de l'allocation aux adultes handicapés | 3 718 | - | 3 718 | 695 | 3 546 | 693 |
| Familles percevant l'allocation de parent isolé longue (hors grossesse) | - | - | 325 | 128 | 302 | 81 |
| Bénéficiaires du complément de l'allocation aux adultes handicapés | 153 | - | 153 | 142 | 119 | 114 |
| Bénéficiaires du minimum vieillesse | - | - | 1 556 | 697 | 1 511 | 697 |
| Familles percevant du revenu minimum d'insertion | 4 111 | 965 | 3 330 | 1 102 | 3 379 | 1 098 |
| Familles percevant de l'allocation d'éducation spéciale | 315 | 109 | 173 | 60 | 125 | 48 |
| Familles percevant des allocations familiales | 10 507 | 4 473 | 10 507 | 4 473 | 10 358 | 4 460 |
| Familles percevant de l'allocation de garde d'enfant à domicile | 134 | 61 | 134 | 61 | 113 | 56 |
| Bénéficiaires de l'allocation parentale d'éducation | 2 761 | - | 2 761 | 537 | 2 722 | 537 |
| Familles percevant l'allocation pour jeune enfant courte | 782 | 413 | 782 | 413 | 721 | 402 |
| Familles percevant l'allocation pour jeune enfant longue | 1 824 | 1 023 | 1 824 | 1 023 | 1 711 | 971 |
| Familles percevant l'allocation de rentrée scolaire | 1 299 | 3 012 | 1 299 | 3 012 | 1 155 | 2 748 |
| Familles percevant l'allocation de soutien familial | 778 | 553 | 203 | 145 | 86 | 64 |
| Familles percevant le complément familial | 1 502 | 926 | 1 502 | 926 | 1 489 | 924 |
| Familles percevant d'une prestation logement | 11 944 | 6 140 | 11 944 | 6 390 | 11 910 | 6 787 |

*Source : les deux premières colonnes du tableau sont alimentées par les statistiques nationales de la CNAF (2001) pour l'année 2000 (les masses financières sont en droits constatés, elles comprennent la contribution au remboursement de la dette sociale). Les deux colonnes suivantes correspondent à des calculs réalisés par François Legendre, Jean-Paul Lorgnet et Florence Thibault – bureau des Prévisions de la CNAF – à partir des données de la CNAF (2001) ou de la Direction de la recherche, des études, de l'évaluation et des statistiques du ministère de l'Emploi et de la Solidarité (DREES) pour le minimum vieillesse. Les deux dernières colonnes sont issues du modèle MYRIADE à partir des calculs des auteurs pour l'année 2000. Les quatre premières colonnes portent sur le champ des ménages de Métropole. Les données issues de MYRIADE se réfèrent au champ des ménages ordinaires de Métropole. Pour le minimum vieillesse, les auteurs disposaient des chiffres de 1999 fourni par la DREES. La masse financière et le nombre de bénéficiaires de 2000 ont été déterminé en appliquant à ces chiffres le taux de croissance moyen de la période 1994-1999 (– 4.5 %).*

Le nombre d'individus bénéficiaires de l'allocation aux adultes handicapés et son complément, retenus comme cible dans MYRIADE, est corrigé afin de prendre en compte l'existence de ce droit multiple au sein des familles (les masses financières ne sont pas modifiées). Il en est de même pour l'allocation parentale d'éducation (dans une famille, le père et la mère peuvent bénéficier simultanément de cette allocation) et le minimum vieillesse (tableau 2, colonnes 3 et 4).

L'enquête Revenus fiscaux qui sert de base pour MYRIADE ne contient pas toute l'information nécessaire à l'imputation de certaines prestations sociales. On a déjà mentionné ce problème pour l'allocation d'éducation spéciale, à l'allocation de soutien familial et à l'allocation de parent isolé. En outre, le choix d'un pas annuel retenu pour développer le modèle oblige à calculer des effectifs cibles différents de ceux des statistiques nationales pour les prestations reposant sur une condition de ressources trimestrielles (tableau 2, colonnes 3 et 4) : RMI et allocation de parent isolé.

## Le calcul des prestations : des résultats légèrement contrastés

Les résultats du calcul des prestations par MYRIADE sont légèrement contrastés (tableau 2, colonnes 5 et 6). Les prestations qui ont fait l'objet d'un étalonnage – RMI, allocation aux adultes handicapés, minimum vieillesse et allocation parentale d'éducation – sont parfaitement simulées. Les allocations familiales sont tout aussi bien traitées en raison du contrôle du type de famille et de la pyramide des âges dans la partie amont de MYRIADE. De même, l'allocation pour jeune enfant, le complément familial et l'allocation de rentrée scolaire donnent des résultats très satisfaisants.

Comme on pouvait s'y attendre, l'allocation d'éducation spéciale, l'allocation de parent isolé et l'allocation de soutien familial sont mal simulées. Il en va de même du complément de l'allocation aux adultes handicapés et de l'allocation de garde d'enfant à domicile. Le manque de données dans l'enquête ne permet pas de couvrir correctement le champ de ces prestations. Enfin, si la masse des allocations logements est celle qui était visée, le nombre de bénéficiaires est surestimé de 5 %.

## Des difficultés pour valider le module de l'impôt sur le revenu

Il n'est pas possible de valider le module de l'impôt sur le revenu comme on a pu valider les prestations familiales et les minima sociaux. En effet, les données pouvant servir au cadrage ne sont disponibles que pour l'année de l'enquête. Dans le cas présent, les chiffres connus sont ceux de 1997. L'année de base du modèle est 2000. Les données de la DGI relatives à l'année 2000 ne seront publiées qu'en 2003. Ce décalage s'explique par la prescription, puisque durant trois années les impositions peuvent être corrigées.

Par ailleurs, l'enquête est représentative des ménages ordinaires tandis que les données de la DGI couvrent tous les foyers fiscaux. MYRIADE n'est pas sur le même champ car il ignore le nombre et les revenus des foyers fiscaux situés dans des ménages non ordinaires. Enfin, une dernière difficulté réside dans la complexité de la législation fiscale. Pour des raisons d'efficacité, il n'a pas été possible de garder toutes les déductions et réductions prévues par le législateur. Ces choix sont à l'origine de discordances entre les chiffres de la DGI et les résultats de MYRIADE.

Deux méthodes servent à valider les résultats du calcul de l'impôt. La première consiste à évaluer pour chaque foyer fiscal la différence entre l'impôt calculé tel qu'il résulte des choix opérés dans MYRIADE et le montant connu de l'impôt réellement acquitté par le ménage. Ce travail au cas par cas valide le barème. Certains écarts observés portent sur la composition des foyers fiscaux. Dans certains foyers, il existe des personnes

rattachées vivant hors du ménage. Ce cas est traité dans un autre module de MYRIADE, il n'affecte aucunement le calcul de l'impôt. D'autres écarts constatés concernent les déductions et réductions non prises en compte par MYRIADE : là non plus le calcul n'est pas affecté.

La deuxième méthode permet de valider la masse de l'impôt calculé. Elle consiste à comparer le calcul de l'impôt obtenu dans le modèle pour les revenus de 1997 avec les données de la DGI. L'écart est inférieur à 0,5 %. Il est le résultat de deux effets jouant en sens inverse : le champ restreint aux ménages ordinaires dans MYRIADE conduit à sous-évaluer la masse de l'impôt, les déductions et réductions non retenues en totalité pour le recalcul de l'impôt dans le modèle impactent dans le sens d'une surestimation.

## L'apport de MYRIADE

La version statique de MYRIADE est complémentaire des autres sources statistiques disponibles à la CNAF en matière d'études, de chiffrage et de prévisions. Elle dispose d'une information moins précise sur les bénéficiaires de prestations sociales et familiales mais son apport est caractérisé par trois effets qu'elle appréhende particulièrement bien, à la différence des autres outils existants : il s'agit des effets de débordement, de composition et de localisation.

L'effet de débordement a pour origine l'élargissement de la population des bénéficiaires à la suite d'une modification d'un critère d'éligibilité. Il est difficile à appréhender soit parce que les nouveaux éligibles ne sont pas connus des CAF, soit parce que l'information nécessaire, pour des individus déjà allocataires, n'est pas disponible. Il en va ainsi des revenus pour les allocataires ne bénéficiant pas de prestations sous conditions de ressources.

Par exemple, MYRIADE a été sollicité pour chiffrer le nombre de familles comprenant des grands enfants, âgés de 20 ans à 24 ans. Ces enfants ne figurent pas dans les fichiers de la CNAF, car ils ne sont comptés à charge pour aucune des prestations servies par les CAF. MYRIADE, de son côté, ne connaît pas cette limite. Ces enfants et leurs parents sont regroupés au sein d'une même famille. Il faut pour cela que les enfants respectent deux conditions : ils ne doivent pas être autonomes financièrement et ne doivent pas être bénéficiaires eux-mêmes d'une prestation. De même, le modèle a pu chiffrer l'impact d'un relèvement du plafond de ressources de l'allocation pour jeune enfant à la suite d'une demande de la Cour des comptes.

L'effet de composition est lié à l'accumulation, pour un même individu ou une même famille, des prélèvements obligatoires et des transferts sociaux. Pour résoudre diverses questions sociales, les pouvoirs publics peuvent modifier les législations fiscales ou sociales. Il est aussi possible d'intervenir à la fois sur les deux législations. Pour rendre compte de telles mesures, il faut pouvoir reconstituer à la fois les prélèvements fiscaux mais aussi les transferts sociaux pour un même individu. Par exemple, il est difficile d'établir le bilan redistributif comparé des différentes aides à la garde des jeunes enfants : ces dernières prennent la forme à la fois d'une subvention directe et d'une réduction d'impôt. Pour expliciter cet enchevêtrement, il faut simuler l'ensemble de la réglementation. La modularité de MYRIADE permet de résoudre le problème. A l'aide du modèle, il est possible d'évaluer chacune des mesures mais aussi d'apprécier l'effet cumulé des mesures fiscales et sociales.

Le troisième effet est l'effet de localisation. Par exemple, on souhaite créer une aide spécifique aux bénéficiaires de l'allocation de parent isolé pour se loger, aide qui dépendrait des ressources de l'ensemble des personnes vivant dans le lo-

gement. Cette mesure ne peut pas être chiffrée à partir des fichiers de la CNAF, car si l'on connaît les ressources de la personne percevant l'allocation de parent isolé, on ignore la composition du ménage et *a fortiori* les revenus globaux du ménage. En revanche, MYRIADE est en mesure de reconstituer l'environnement de tout individu. Plus généralement, le modèle retrace le passage des revenus individuels aux ressources disponibles de la famille ou du ménage. Il lui est possible, par exemple, d'évaluer les phénomènes de polarisation entre les familles.

Pour résumer, l'apport de MYRIADE réside en trois points. Tout d'abord, le modèle est représentatif de l'ensemble des ménages ordinaires de la population de Métropole. Ensuite, il situe chaque individu dans son environnement, que ce soit son ménage, son foyer fiscal ou encore sa famille. Enfin, MYRIADE simule la plupart des instruments du système de redistribution en France. Toutefois, certains transferts ne sont pas simulés tels que les aides sociales locales, les dépenses publiques d'éducation et les remboursements d'assurance maladie. Il en est de même pour certains prélèvements : la TVA, la taxe intérieure sur les produits pétroliers et les taxes sur l'alcool et le tabac.

Cet article a poursuivi un double objectif : d'une part, présenter le modèle de microsimulation de la CNAF et, d'autre part, montrer ses potentialités à l'aide d'exemples illustratifs. Il a semblé nécessaire de détailler les choix structurants arrêtés pour développer le modèle et les principales hypothèses – contestables – retenues. Les trois autres articles de ce dossier de ***Recherches et Prévisions*** (6) sont démonstratifs des possibilités ouvertes par la version statique du modèle : dresser un portrait de certaines inégalités sociales, comprendre les principales propriétés du système de redistribution français, chiffrer des mesures nouvelles et en montrer les conséquences redistributives.

(6) Voir les articles *Les modèles socio-économiques de microsimulation. Panorama et état des lieux pour la France* (p. 11), *La redistribution au bénéfice des familles : l'apport de MYRIADE (p. 51)* et *Explorations autour de l'allocation familiale unique. Substituer une seule prestation aux prestations d'entretien des enfants existantes (p. 71)* dans ce numéro de ***Recherches et Prévisions***.

## Références bibliographiques


Accardo J. et Fall M., « *La mesure des revenus dans les enquêtes ménages et fiabilité des indicateurs d'inégalités dérivés* », Document de travail INSEE, F 9602, 1996.

Amrouni I. et Thibault F., *L'échantillon national des allocataires : un outil pour des études et des simulations*, ***Recherches et Prévisions***, 2001, n° 65.

Atkinson A. B., Bourguignon F. et Chiappori P. A. , *What do we learn about tax reform from international comparison ? France and Britain*, ***European Economic Review***, 1998, vol. 32.

Blanchet D., *La microsimulation appliquée à l'analyse des politiques sociales*, ***Economie et Statistique***, 1998, n° 315.

Blanchet D. et Chanut J.-M., *Les retraites individuelles à long terme : une projection par microsimulation*, ***Economie et Statistique***, 1998, n° 315.

Bonnet C. et Mahieu R., « Microsimulation techniques applied to intergenerational transfers-Pension in a dynamic framework : the case of France », Document de travail INSEE, G 9906, 1999.

Bourguignon F., O'Donoghue C., Sastre-Descals J., Spadaro A. et Utili F., *Eur 3 : a Prototype European Tax-Benefit Mode*, **DAE Working Papers**, Cambridge, University of Cambridge, 1997, n° 9323.

David M.-G., Lhommeau B. et Starzec C., « Le Modèle de microsimulation INES INSEE - Etudes Sociales : description et exemples d'utilisation », Document de travail INSEE, F 9902, 1999.

Division Redistribution et Politiques Sociales, « Le modèle de microsimulation dynamique DESTINIE », Document de travail INSEE, G 9913, 1999.

Grignon M. et Pennec S., « Simuler les effets et prévoir les contraintes de la politique familiale en France », communication à la journée « Application des méthodes de microsimulation », Paris, 1996.

Imhoff Van E. et Post W., *Méthodes de microsimulation pour des projections de population*, ***Population***, 1997, vol. 53, n° 4.

Jeandidier B., *La spécificité des politiques familiales en Europe. Une application menée à l'aide de microsimulations*, ***Recherches et Prévisions***, 1997, n° 48.

Legendre F., « Les modèles socio-économiques de microsimulation : panorama et état des lieux pour la France », Document de travail 99-05, Conseil supérieur de l'emploi, des revenus et des coûts, 1999.

Legendre F., « Un prototype d'un modèle de micro-simulation destiné à l'étude des politiques fiscales et sociales », note interne, EPS-ADEPS, Université Nancy-II et CNRS, 1999.

Merz J., *Microsimulation – a survey of principles, developments and applications*, ***International Journal of Forecasting***, 1991, vol. 7.

Orcutt G. H., *A new type of socio-economic systems*, ***The Review of Economics and Statistics***, 1957, vol. 58.

Stroustrup B., ***The C++ programming language***, Addison-Wesley, 1997.

Sutherland H., *Les modèles statiques de microsimulation en Europe dans les années 90*, ***Economie et Statistique***, 1998, n° 315.


*Les documents de travail de l'INSEE sont disponibles sur le site de l'institut : http://www.insee.fr*