



RÉPUBLIQUE
FRANÇAISE

*Liberté
Égalité
Fraternité*



LES DOSSIERS DE LA DREES

N° 87 • décembre 2021

Le modèle de micro-simulation Ines-Omar 2017

Note méthodologique détaillée

Mathieu FOUQUET (DREES)



STATISTIQUE
PUBLIQUE

La DREES fait partie
du Service statistique
public piloté par l'Insee.

Le modèle de micro-simulation Ines-Omar 2017

Note méthodologique détaillée

Mathieu FOUQUET (DREES)

Remerciements : Mathilde GAINI, Catherine POLLAK, Noémie COURTEJOIE, Raphaële ADJERAD (DREES)

SOMMAIRE

Table des matières

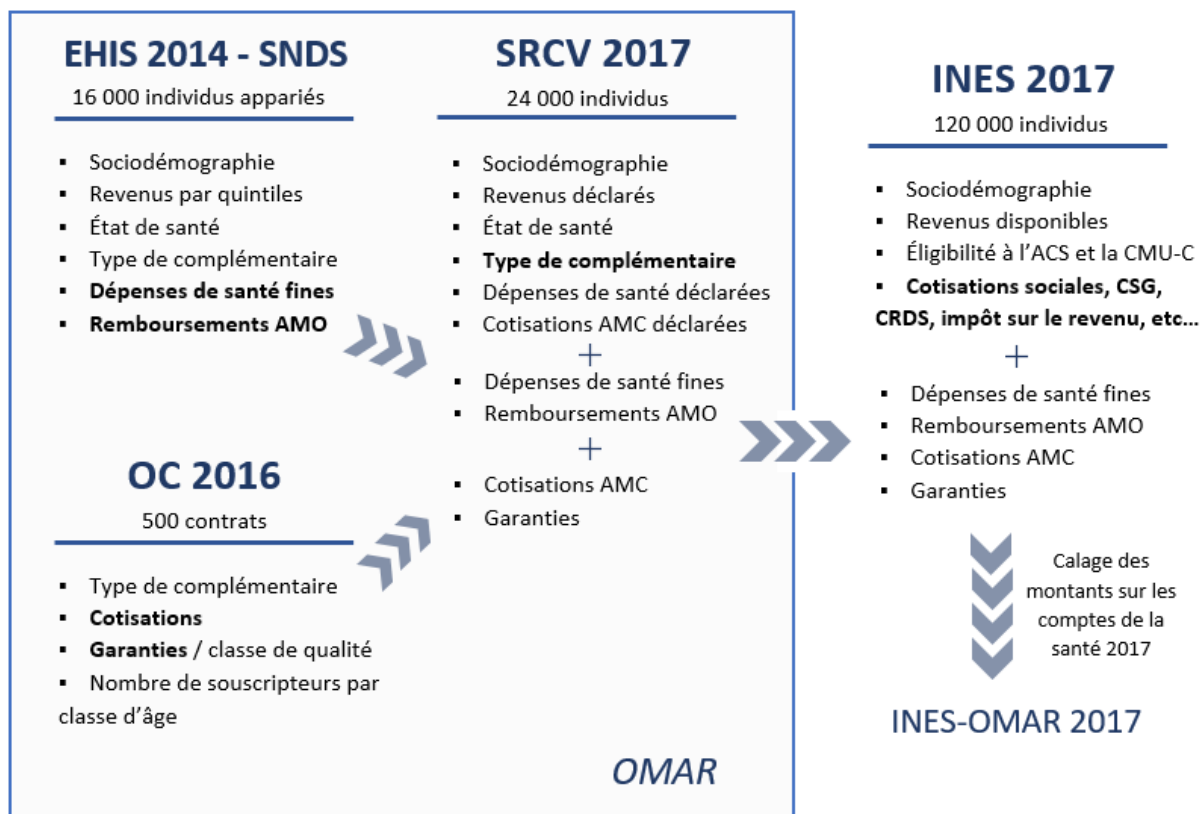
■ INTRODUCTION	2
■ LES RETRAITEMENTS SUR LA COMPLÉMENTAIRE SANTÉ.....	4
Attribution d'une couverture dans SRCV 2017	4
Contrats d'entreprise ou individuels.....	5
■ LES DÉPENSES DE SANTÉ.....	7
■ PSEUDO-APPARIEMENT EHS-SRCV	9
■ ATTRIBUTION D'UN CONTRAT DE COMPLÉMENTAIRE SANTÉ.....	11
Retraitement du type de contrat	11
Imputation des contrats.....	11
■ CALCUL DES COTISATIONS DE COMPLÉMENTAIRE SANTÉ.....	13
Retraitements des informations de cotisation dans l'enquête OC	13
Simulation fine des cotisations	13
Vérifications	14
■ ESTIMATION DES REMBOURSEMENTS DE COMPLÉMENTAIRE SANTÉ.....	16
Modes d'expression des garanties	16
Niveau de détail des simulations.....	17
■ LE FINANCEMENT DE L'ASSURANCE MALADIE OBLIGATOIRE.....	18
Pseudo-appariement à Ines 2017	18
Calages macro-économiques	19
Distribuer le financement de l'AMO	21
■ POUR EN SAVOIR PLUS.....	23
Annexe 1. Questionnaire couverture complémentaire de SRCV	24

INTRODUCTION

Le modèle Omar (outil de micro-simulation pour l'analyse des restes à charge) permet d'étudier la redistribution opérée par le système d'assurance maladie du point de vue des ménages. Il fournit des informations sur les dépenses de santé, les remboursements versés respectivement par l'assurance maladie publique et par les complémentaires santé, sur les cotisations dont doivent s'acquitter les souscripteurs à une complémentaire et sur ce qui demeure à la charge des ménages après remboursements complémentaires. Ces données peuvent être déclinées sur diverses sous-populations selon le type de couverture complémentaire, le niveau de vie, l'âge, le fait ou non d'être un patient en affection de longue durée (ALD), l'état de santé déclaré, le statut d'activité. Lorsque le modèle Omar est pseudo-apparié au modèle Ines, on dispose d'éléments additionnels sur la participation des ménages au financement de l'assurance maladie obligatoire (AMO). Il devient alors possible de calculer le poids de l'ensemble des dépenses de santé et de leur financement dans le revenu des ménages et d'étudier la redistribution opérée par le système d'assurance maladie dans son ensemble.

Ce document présente la méthode de la version 2017 du modèle Ines-Omar (*schéma 1*). Contrairement aux précédentes versions du modèle, basées sur l'Enquête santé et protection sociale (ESPS) de l'Institut de recherche et documentation en économie de la santé (Irdes), le cœur de cette nouvelle édition du modèle est l'enquête Statistique sur les ressources et les conditions de vie (SRCV) 2017 produite par l'Institut national de la statistique et des études économiques (Insee). Il s'agit de la première enquête en population générale permettant de connaître l'état de couverture de la population par une complémentaire santé à la suite de la généralisation de la complémentaire d'entreprise en 2016. En effet, l'appariement entre l'Enquête santé européenne (EHIS) 2019¹ et le Système national des données de santé (SNDS) n'était pas encore mobilisable au moment de la mise à jour du modèle.

Schéma 1 • Schéma des éléments constitutifs du modèle Ines-Omar 2017



Nous disposons dans SRCV 2017 d'informations sur les caractéristiques sociodémographiques de plus de 24 000 individus, ainsi que sur leur niveau de vie, le type de leur couverture complémentaire santé, leur état de santé perçu, leurs dépenses de santé déclarées et leur cotisation de complémentaire santé déclarée. Néanmoins, la connaissance du statut de couverture de chaque individu nécessite un certain nombre de retraitements (voir partie II). En outre, SRCV 2017 n'étant pas appariée au SNDS, elle ne permet pas de connaître les dépenses de santé avec un niveau de finesse suffisant. En effet, seules sont disponibles des informations déclarées sur les visites récentes chez le médecin, chez le dentiste ou à l'hôpital, et sur les

¹ L'Enquête santé européenne a remplacé l'Enquête santé et protection sociale (ESPS), menée par l'Irdes, à partir de l'édition 2014.

consommations récentes de médicaments. Ceci permet de se faire une idée du niveau de dépenses de santé d'un individu, mais ne permet nullement d'effectuer des calculs précis sur les montants de dépenses et de remboursements impliqués.

La connaissance fine des dépenses de santé provient de l'appariement EHIS-ESPS 2014 – SNDS, produit conjointement par l'Irdes et la Direction de la recherche, des études, de l'évaluation et des statistiques (DREES). Ceci permet de disposer pour plus de 16 000 individus des mêmes données individuelles que celles présentes dans SRCV 2017, associées à des données exhaustives et très détaillées sur les dépenses de santé et les remboursements perçus au titre de l'AMO (voir partie III). Les individus de SRCV sont ainsi pseudo-appariés à des individus de EHIS afin que des données de santé leur soient imputées (voir partie IV).

L'enquête de la DREES auprès des organismes complémentaires sur leurs contrats de complémentaire santé les plus sous-crits (enquête OC) donne de nombreux éléments sur les contrats individuels et collectifs : nombre de souscripteurs par classe d'âge ; cotisations en fonction de l'âge, du nombre de bénéficiaires, du revenu et du lieu de résidence ; niveaux de garantie pour chaque poste de soin. Nous attribuons un contrat tiré de l'enquête à chaque ouvrant droit de contrat présent dans SRCV (voir partie V) et nous recalculons finement sa cotisation (voir partie VI), ainsi que les remboursements complémentaires perçus par lui et ses ayants droit (voir partie VII). Nous aboutissons ainsi au modèle Omar 2017.

Le modèle Ines, développé conjointement par l'Insee, la DREES et la Caisse nationale des allocations familiales (CNAF), permet de simuler la législation socio-fiscale (Fredon, Sicsic 2020). Il s'appuie sur l'Enquête revenus fiscaux et sociaux (ERFS) pour connaître précisément l'ensemble des revenus perçus par près de 120 000 individus et simule les prélèvements sociaux et fiscaux dont ils doivent s'acquitter. Un module complémentaire a été développé pour simuler les formes de taxation indirecte : TVA, taxes sur l'alcool et le tabac (André, *et al.*, 2016). Les individus de la version 2017 du modèle Ines sont pseudo-appariés à des individus du modèle Omar 2017 sur la base de l'ensemble des caractéristiques socio-démographiques connues. À la suite de ce pseudo-appariement, il devient possible d'estimer la contribution de différentes sous-populations au financement de l'AMO (voir partie VIII).

Enfin les montants de dépenses de santé sont calés sur le niveau macro-économique de dépense établi par les comptes nationaux de la santé (CNS) 2017 sur le champ de la dépense remboursable et individualisable. La distribution de la dépense entre les ménages ainsi que les parts remboursées par l'AMO et l'assurance maladie complémentaire (AMC) sont celles estimées par le modèle (voir partie VIII).

Le schéma 1 résume les différents éléments constitutifs du modèle et leur articulation, ainsi que la provenance des données du modèle finalisé. Étant donné que la construction du modèle consiste en l'imputation de données issues d'Omar aux individus d'Ines, et non l'inverse, il est à noter que les principales caractéristiques sociodémographiques des individus du modèle Ines-Omar 2017 sont ceux du modèle Ines, notamment l'âge, le sexe, l'occupation et les revenus. Le champ des dépenses de santé considérées est celui de l'ensemble des dépenses individualisables et remboursables par l'AMO en France entière, hors dépenses exécutées en milieu médico-social. Il n'inclut pas les indemnités journalières ni les rentes pour accidents du travail et maladies professionnelles. Le modèle est représentatif de l'ensemble des ménages ordinaires de France métropolitaine. Étant donné que ce champ diffère légèrement de celui considéré pour les dépenses de santé, nous faisons implicitement l'hypothèse que les montants moyens de dépenses de santé sont semblables pour les ménages dans et hors champ, notamment entre la métropole et les départements d'outre-mer.

■ LES RETRAITEMENTS SUR LA COMPLÉMENTAIRE SANTÉ

Les questions portant sur la complémentaire santé n'étant qu'un module complémentaire de SRCV 2017, celles-ci sont limitées et de nombreux retraitements sont nécessaires avant toute exploitation. En effet il est impératif :

- d'attribuer les couvertures aux ayants droit au sein d'un ménage. En effet, bien que l'ensemble des personnes de 16 ans ou plus soient interrogées sur leur couverture complémentaire santé, le lien entre ayant droit et ouvrant droit n'est pas établi explicitement. De plus, le statut de couverture des enfants de moins de 16 ans est inconnu ;
- de corriger les réponses sur la nature du contrat (individuel ou d'entreprise), car un certain nombre de personnes ne connaissent pas ou mal le type de contrat dont ils bénéficient, conduisant à des situations incohérentes. Ce retraitement est également nécessaire pour utiliser l'EHIS 2014.

Les seuls ménages supprimés sont ceux composés uniquement de non-répondants au module complémentaire santé. On passe ainsi de 25 401 individus à 24 520 personnes dans la base finale.

Attribution d'une couverture dans SRCV 2017

Il est nécessaire d'établir pour chaque individu d'un ménage son état de couverture précis : type de contrat, ayant droit ou ouvrant droit, personnes du ménage rattachées au même contrat, cotisation versée pour ce contrat. Cependant, la couverture complémentaire au sein des ménages est mal connue :

- les enfants ne font pas partie du questionnaire. Ils sont donc seulement comptabilisés indirectement par leur ouvrant droit via la variable qui recense le nombre d'ayants droit couverts par lui (variable CombPers) ;
- un certain nombre d'ayants droit adultes n'ont pas répondu au questionnaire, ou seulement partiellement ;
- le lien explicite entre l'ouvrant droit et ses ayants droit n'est pas établi par le questionnaire. Il doit donc être établi au mieux en utilisant toutes les informations connues sur les contrats.

Ainsi, sans retraitement, seulement 77 % de la population serait considérée comme couverte par une complémentaire santé et, pour beaucoup, il manquerait la connaissance des caractéristiques du contrat. Après retraitement, nous estimons que plus de 96 % de la population est aujourd'hui couverte. Les résultats issus de ce retraitement ont été publiés dans Fouquet, 2020.

L'attribution de la couverture maladie universelle complémentaire (CMU-C)

Lorsqu'au moins une personne du ménage s'est déclarée bénéficiaire de la CMU-C, et si cette personne est considérée comme un adulte du ménage, la CMU-C est affectée aux autres membres du ménage car elle est en principe attribuée à l'ensemble des membres d'un ménage. L'ouvrant droit retenu est la personne la plus âgée ayant déclaré bénéficiaire de la CMU-C. Sont considérés comme n'étant pas bénéficiaires : les personnes d'une génération antérieure à celle de l'ouvrant droit et les enfants de l'ouvrant droit ayant plus de 25 ans, sauf s'ils ont déclaré explicitement bénéficiaire de la CMU-C. Tous les autres membres du foyer sont considérés bénéficiaires de la CMU-C.

L'identification des bénéficiaires de la CMU-C par le questionnaire complémentaire de SRCV est assez limitée, n'incorporant aucun contrôle de cohérence pour s'assurer que l'enquête ne s'est pas trompé. Il est donc probable que des personnes bénéficiaires de la CMU-C aient déclaré être couvertes par un contrat individuel privé et soient par conséquent considérées comme telles malgré les retraitements effectués. Le taux de couverture par la CMU-C estimé à 6,6 % après retraitement dans SRCV est ainsi inférieur au taux de 8,1 % estimé par le Fonds CMU à la période où l'enquête SRCV était menée².

L'aide à la complémentaire santé (ACS) complète le dispositif CMU-C en soutenant les ménages légèrement au-dessus du seuil d'éligibilité à la CMU-C dans le financement de leur complémentaire santé. Le module complémentaire de SRCV 2017 ne permet pas de savoir qui est couvert par l'ACS. Le recours à ce dispositif est donc attribué aléatoirement à des individus correspondant à la bonne tranche de revenus, de façon à respecter les effectifs de recourants par tranche d'âge observés dans le SNDS.

² Cette sous-estimation ne perturbe pas le fonctionnement du modèle Ines-Omar 2017. En effet, les bénéficiaires de la CMU-C dans le modèle finalisé sont tirés au sort parmi les ménages d'Ines éligibles au dispositif, de manière à avoir le nombre exact de bénéficiaires de la CMU-C tel qu'estimé par le Fonds CMU. La connaissance de la CMU-C dans SRCV n'est mobilisée qu'afin de pseudo-apparier entre eux les bénéficiaires des différentes sources.

Le nombre de personnes couvertes par chaque ouvrant droit

L'une des principales variables utilisées pour le retraitement est CombPers, qui recense pour tout ouvrant droit le nombre de personnes rattachées à son contrat dans le ménage, y compris lui-même. Deux types d'erreurs ont pu être commises par les répondants à l'enquête :

- les ouvrants droit peuvent ne pas s'être inclus eux-mêmes dans le décompte. Pour corriger au mieux cela, lorsqu'une seule personne dans le foyer n'est pas couverte, ou lorsqu'une seule personne parmi celles déclarées couvertes manque de couverture, l'ouvrant droit est considéré comme s'étant trompé et la variable CombPers est augmentée de 1. Cette correction concerne 269 enquêtés. À l'inverse, si plusieurs personnes manquent de couverture, il est considéré que l'ouvrant droit ne s'est pas trompé et que ces personnes ne sont pas couvertes ;
- les ouvrants droit peuvent ne pas avoir compris la question et ne s'être comptés qu'eux-mêmes malgré l'existence d'ayants droit. Si aucun ouvrant droit du foyer n'a déclaré plus d'une personne rattachée à son contrat, mais que le ménage compte au moins une personne qui se déclare rattachée au contrat d'un autre membre du ménage, il est considéré que l'un des ouvrants droit n'a pas compris la question et qu'il couvre potentiellement tous ceux déclarés couverts. Le cas échéant, étant donné que le statut de couverture des enfants de moins de 16 ans est inconnu et que suite à l'erreur de l'enquêté ils n'ont pas été comptabilisés dans CombPers, leur statut de couverture est tiré au sort en respectant le taux moyen de couverture. Cette correction concerne 120 ménages.

Logique de couverture

Une fois la variable CombPers corrigée, l'enjeu est tout d'abord d'identifier tous les membres du ménage couverts, même s'ils sont enfants ou non-répondants ; ensuite de rattacher chaque ayant droit à son ouvrant droit en fonction des informations disponibles sur son contrat. L'algorithme mis en œuvre est le suivant :

- la variable CombPers indique le nombre total de « couvertures » qui doivent être attribuées dans le ménage. Si ce nombre est égal au nombre d'ouvrants droit, il n'y a alors aucun retraitement à effectuer ;
- sinon, il faut rattacher chaque personne déclarée couverte à son ouvrant droit : il est recherché au sein des différents contrats souscrits par le ménage celui correspondant le mieux aux informations fournies par l'ayant droit (individuel ou d'entreprise, montant des cotisations, satisfaction vis-à-vis des remboursements ou, à défaut de tout cela, tirage au sort) ;
- le nombre de couvertures du ménage peut être supérieur au nombre de personnes déclarées couvertes si le ménage contient des enfants, des adultes non répondants, ou si un membre du ménage a une mauvaise connaissance de sa situation de couverture. Pour chaque contrat, les bénéficiaires sont alors tirés au sort parmi les membres du ménage potentiellement couverts.
- Enfin, il se peut que certains membres du foyer soient déclarés couverts mais qu'aucun ouvrant droit ne se soit déclaré. Un des adultes couverts est alors tiré au sort comme étant l'ouvrant droit, et les enfants se voient appliquer le taux de couverture moyen.

Contrats d'entreprise ou individuels

Les questions adressées aux ménages dans EHIS et dans SRCV permettent d'identifier si la complémentaire santé est un contrat d'entreprise. Les contrats collectifs qui ne sont pas des contrats d'entreprise ne peuvent cependant pas être distingués des contrats individuels et sont donc considérés comme tels.

Il est fréquent que des personnes bénéficiant de contrats individuels, notamment les fonctionnaires, pensent être couvertes par une complémentaire d'entreprise car elles ont souscrit auprès d'une mutuelle spécialisée dans la couverture santé des agents d'un ministère spécifique. Ces erreurs sont corrigées à l'aide du statut d'activité et de la catégorie sociale de l'ouvrant droit, pour lui ainsi que pour tous ses ayants droit. Bien que les variables soient différentes entre SRCV et EHIS, la logique suivie est identique : parmi les personnes couvertes par une complémentaire privée, sont considérés comme bénéficiaires de contrats individuels les fonctionnaires, les indépendants, les personnes sans emploi depuis plus d'un an et les retraités.

Le tableau 1 récapitule les effets des différents retraitements sur la situation de couverture individuelle dans SRCV. Les complémentaires d'entreprise ayant en moyenne davantage de bénéficiaires par ouvrant droit, leur proportion a augmenté plus fortement que celle des contrats individuels lors de la première étape de retraitement.

Tableau 1 • Répartition des situations de couverture dans la population avant et après retraitements de l'enquête SRCV 2017

	Avant retraitements	Après corrections du statut de couverture	Après tous retraitements
Complémentaires d'entreprise	43,2	59,3	38,1
Contrats individuels	29,7	30,2	51,4
CMU-C	3,9	6,6	6,6
Non couverts	23,2	3,9	3,9

Lecture > Avant retraitements, 43,2 % des individus de l'enquête étaient considérés couverts par une complémentaire d'entreprise. Ils étaient 59,3 % après la première étape de retraitement, et plus que 38,1 % à l'issue de l'ensemble des retraitements.

Champ > Population française métropolitaine résidant en ménage ordinaire.

Source > SRCV 2017. Traitements DREES.

■ LES DÉPENSES DE SANTÉ

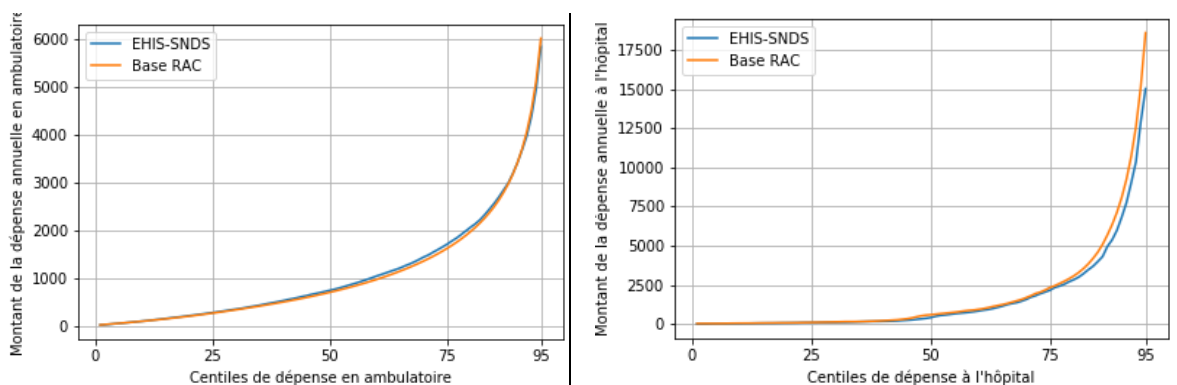
Les dépenses de santé du modèle Ines-Omar 2017 sont issues de l'appariement de l'EHIS 2014 au SNDS. La jointure EHIS-SNDS n'est possible que pour 16 757 individus de l'enquête. Un recalage des poids individuels selon l'âge, le sexe et le niveau de vie est effectué pour que ce sous-échantillon demeure représentatif de la population de France métropolitaine.

Le SNDS contient des informations sur les soins remboursés en ville ou en établissements de santé. Il est possible de calculer le montant des dépenses de santé, la part remboursée par l'AMO, la part de liberté tarifaire ou encore les participations forfaitaires acquittées. L'appariement exhaustif des données sur les dépenses de santé en ville et en hôpital privé, issues du DCIRS (datamart de consommation inter-régime simplifié) et des dépenses de santé en hôpital public, issues du PMSI (Programme de médicalisation des systèmes d'information), dans le cadre de la base RAC (reste à charge) a permis de fiabiliser le calcul de ces indicateurs (Adjerad, Courtejoie, 2021). Le champ retenu pour les postes de soins est le plus exhaustif possible. Il couvre les dépenses remboursables et présentées au remboursement (qu'il y ait eu avance de frais ou non), réalisées en 2014 en ville, au cours d'un séjour à l'hôpital public ou en établissement de santé privé, y compris les soins de suite et de réadaptation (SSR), la psychiatrie (PSY) et l'hospitalisation à domicile (HAD). Comme le précise l'encadré méthodologique décrivant la base RAC dans Adjerad, Courtejoie, 2021, « les dépenses non directement individualisables sont exclues de l'analyse. Il s'agit notamment de celles associées à la rémunération des médecins libéraux sur objectifs de santé publique, qui ne peuvent pas être directement affectées à un bénéficiaire spécifique. Pour les patients du champ médico-social, qui incluent notamment les personnes handicapées ou âgées en établissements spécialisés, les dépenses incluses dans les forfaits de soins de ces établissements médico-sociaux ne sont pas prises en compte, car elles ne sont que partiellement renseignées dans le SNDS. » Enfin, on ne tient pas compte des dépenses non prises en charge par l'AMO, parce qu'elles ne figurent pas de manière exhaustive dans le SNDS, telles que les suppléments pour confort personnel. On aboutit ainsi au champ des dépenses de santé présentées au remboursement, individualisables et remboursables en ville et à l'hôpital, hors médico-social.

Par ailleurs, pour les séjours hospitaliers, il est nécessaire de recalculer certains montants à partir des informations disponibles, comme précisé dans les fiches de la documentation collaborative du SNDS, disponibles ici : <https://documentation-snds.health-data-hub.fr>. Ainsi, le taux de remboursement est recalculé en fonction de la nature de l'assurance (régime de base, régime Alsace-Moselle, maternité, accident du travail, prévention maladie) et de l'existence éventuelle d'une exonération, tandis que le forfait journalier est recalculé selon la durée du séjour, la nature complète ou partielle de l'hospitalisation, le mode de sortie et les éventuelles exonérations. Un travail a été effectué afin d'intégrer les dépenses associées aux médicaments et dispositifs de la liste en sus, ainsi que les actes et consultations externes.

Enfin, le regroupement des prestations de soins en postes de soins plus généraux nécessite un travail important de croisement des différentes tables du SNDS et d'appariement aux nomenclatures idoines. Seul ce travail fin permet de distinguer les différentes formes de pharmacie, de séparer les prothèses dentaires de l'orthodontie et des actes de conservation dentaire, ou encore l'optique du matériel médical et des audioprothèses.

Graphique 1 • Distribution par centiles du montant total individuel de dépenses de santé à l'hôpital ainsi qu'en ambulatoire. Comparaison entre la base RAC 2017 et EHIS-SNDS 2014



Note > Les centiles de dépenses sont affichés jusqu'au 95^e inclus pour ne pas écraser la courbe.

Lecture > Le troisième quartile des dépenses les plus élevées en ambulatoire s'élève à 1 622 euros en 2017 d'après la base RAC, et à 1 709 euros en 2014 d'après l'échantillon EHIS-SNDS.

Champ > Ensemble des personnes présentes dans l'appariement EHIS 2014-SNDS. Toutes dépenses en hôpital public, y compris SSR/PSY/HAD et dépenses en sus du groupe homogène de séjour. Toutes dépenses en ambulatoire y compris les actes et consultations externes, à l'exclusion du médico-social.

Sources > Base RAC 2017, EHIS-SNDS 2014. Traitements DREES.

Les programmes développés dans le cadre de la base RAC sont repris et exécutés ici sur l'échantillon EHIS-SNDS 2014. Une documentation plus extensive, ainsi que les programmes de la base RAC sont disponibles sur le Gitlab public de la DREES : https://gitlab.com/DREES_code.

Malgré la faible dimension de la base EHIS-SNDS 2014, le graphique 1 montre que la distribution de la dépense de santé totale par centiles a un profil très proche de celle obtenue sur le SNDS exhaustif, en particulier en ambulatoire. Néanmoins, un écart est observé sur l'hôpital que la hausse des dépenses de santé entre 2014 et 2017 ne suffit pas à interpréter. Il est probable que les personnes ayant les plus fortes dépenses de santé soient sous-représentées dans l'enquête EHIS, cette enquête n'ayant pas pu se faire auprès de personnes hospitalisées.

■ PSEUDO-APPARIEMENT EHS-SRCV

Les données de santé des individus présents dans l'EHIS 2014 sont imputées aux individus de SRCV 2017 par pseudo-appariement. Plusieurs méthodes de pseudo-appariement ont été comparées : *hot-deck*, arbres de régression et forêts aléatoires. Étant donné que les deux enquêtes contiennent essentiellement les mêmes informations et que les questionnaires sont très proches, il est apparu judicieux de recourir à une méthode permettant de mobiliser cette importante quantité de données. Ainsi, le pseudo-appariement s'effectue en s'appuyant sur des variables aussi diverses que l'âge, le sexe, le niveau de vie, le statut ALD, l'état de santé déclaré, le statut d'occupation, le type de couverture complémentaire santé, le niveau d'études, le nombre de personnes du ménage, les consommations de soins déclarées (arrêts de travail, visites chez le médecin, chez le dentiste ou à l'hôpital, consommation de médicaments) ou encore les habitudes de vie.

Arbres de régression

Les arbres de régression sont une méthode d'apprentissage statistique ressemblant fortement au *hot-deck* en ce qu'ils consistent eux aussi à répartir les individus dans des strates homogènes. Cependant les données d'entrée (EHIS 2014-SNDS) sont mobilisées pour déterminer quelles variables sont explicatives de la dépense de santé, comment les subdiviser et comment les hiérarchiser entre elles ; tandis que le *hot-deck* est intégralement paramétré par le statisticien, ce avec moins de souplesse.

L'échantillon de départ est découpé entre un échantillon d'apprentissage sur lequel l'algorithme va « apprendre » et un échantillon de vérification qui permet de mesurer la qualité prédictive de l'arbre obtenu. L'arbre est généré automatiquement par l'algorithme suivant :

- pour chaque variable explicative disponible, trouver le seuil de coupure permettant de scinder au mieux l'échantillon en deux, c'est-à-dire en maximisant l'homogénéité de la variable expliquée, ici le montant total de dépenses de santé de l'individu, au sein de chaque sous-échantillon. Par exemple estimer que c'est l'âge de 72 ans qui sépare le mieux les personnes ayant des dépenses de santé élevées des personnes ayant des dépenses de santé faibles ;
- déterminer quelle variable est la plus explicative de la dépense de santé, puis séparer l'échantillon d'apprentissage en deux sous-groupes de part et d'autre du seuil optimal précédemment trouvé ;
- on applique de nouveau les deux étapes précédentes à chacun des sous-groupes obtenus ;
- l'algorithme s'arrête lorsqu'un sous-groupe atteint une taille trop faible pour être à nouveau subdivisé. L'ensemble des individus du groupe font alors partie d'une même strate. La taille maximale d'une strate est fixée par le statisticien ;
- la subdivision en strates obtenue est appliquée aux données de l'échantillon de vérification afin que le statisticien puisse vérifier les bonnes performances prédictives du modèle et modifier le paramétrage de l'arbre si nécessaire.

Forêts aléatoires

La forêt aléatoire vise à faire des prédictions de meilleure qualité qu'un simple arbre de régression avec la même quantité de données. Au lieu de générer un arbre unique et complexe, on fait pousser une grande quantité d'arbres plus simples. Ces arbres sont générés en tirant au sort avec remise les individus de l'échantillon d'apprentissage et en ne conservant que certaines variables explicatives également tirées au sort. À l'issue, la dépense de santé prédite est la moyenne des dépenses prédites par chaque arbre. En résumé la forêt aléatoire consiste à multiplier les points de vue possibles afin de fournir une interprétation plus robuste des données.

La méthode classique décrite ici permet de prédire pour chaque individu la valeur de la variable expliquée, en l'occurrence son montant de dépenses de santé total. Or notre objectif est d'imputer aux individus de SRCV non pas seulement un montant global mais des données de santé fines décomposées entre postes de soins et entre remboursements, restes à charge, liberté tarifaire, etc... C'est pourquoi il ne s'agit pas de prédire la valeur d'une variable mais d'associer à chaque individu de SRCV un donneur issu de EHS duquel seront tirées l'ensemble des dépenses de santé fines. Ceci nécessite de modifier le fonctionnement de la forêt aléatoire pour en faire un outil de pseudo-appariement plutôt qu'une méthode de prédiction. Pour ce faire, pour un individu de SRCV donné, chacun des arbres de régression calculés pour générer la forêt sélectionne un ensemble d'individus de EHS (une strate) qui semblent bien lui correspondre. Il est ensuite calculé pour chaque individu de EHS combien de fois il a été associé à un même individu de SRCV. Enfin, un donneur est tiré au sort parmi les dix individus de EHS les plus souvent pseudo-appariés au receveur des données, et ce sont les données de santé fines de ce donneur qui sont imputées à l'individu de SRCV considéré.

Les forêts aléatoires réalisées comptent 100 arbres. Chacun de ces arbres est basé sur une sélection de 18 variables explicatives différentes, dont la probabilité de sélection dépend du pouvoir explicatif de la variable. Enfin les feuilles de ces arbres doivent compter au moins 80 individus.

Étant donné les spécificités vis-à-vis du remboursement AMO les concernant, les individus de SRCV présentant une ALD doivent impérativement être pseudo-appariés spécifiquement à des individus de EHIS également en ALD. La méthode d'apprentissage statistique va bien sûr considérer que la présence d'une ALD est un des déterminants majeurs à prendre en compte pour le pseudo-appariement, mais l'existence d'une part d'aléas pourrait conduire quelques bénéficiaires d'ALD à se voir imputer des dépenses de santé de non-bénéficiaires. Il est donc fait le choix de mettre en œuvre un processus de pseudo-appariement spécifique uniquement entre bénéficiaires d'ALD. Il en va de même pour les bénéficiaires de la CMU-C.

■ ATTRIBUTION D'UN CONTRAT DE COMPLÉMENTAIRE SANTÉ

SRCV 2017 contient des éléments sur les contrats de complémentaire santé souscrits par les enquêtés : la nature individuelle/d'entreprise/CMU-C/non couvert, la satisfaction vis-à-vis des remboursements et la prime mensuelle pour l'ensemble des assurés. Mais ces informations sont insuffisantes pour estimer finement les remboursements par l'organisme complémentaire (OC), et il n'est pas certain que les primes déclarées par les enquêtés soient d'une fiabilité suffisante. Pour ces deux raisons il est fait appel à l'enquête OC 2016 sur les contrats les plus souscrits, qui contient de nombreuses informations sur les garanties liées aux contrats ainsi que sur le mode de tarification.

Retraitement du type de contrat

Dans SRCV, les contrats de complémentaire santé sont divisés entre individuels, d'entreprise et CMU-C. Les contrats collectifs qui ne sont pas des contrats d'entreprise sont donc comptés parmi les contrats individuels. Dans l'enquête OC, la distinction se fait entre individuel et collectif. Il est donc nécessaire de reclasser les contrats collectifs non d'entreprise en contrats individuels. Ceci est possible car les organismes complémentaires déclarent si les contrats sont réservés à une certaine catégorie de population.

Le tableau 2 présente les contrats collectifs selon la population à laquelle ils sont réservés. Le score de contrat mentionné dans ce tableau correspond à une méthode d'évaluation du niveau de garantie mise au point par la DREES. Il s'interprète comme un niveau de prise en charge théorique, par les contrats de complémentaire santé, du reste à charge après intervention de l'AMO. Le score varie entre 0 (le contrat n'offre aucune prise en charge du reste à charge sur les prestations de soins considérées) et 1 (le contrat offre une prise en charge complète des restes à charge pour les prestations de soins considérées) [Loiseau, 2020].

Tableau 2 • Statistiques descriptives sur les contrats collectifs selon la population visée

	Retraités, anciens salariés	Fonctionnaires	Salariés du privé	Branche professionnelle	Indépendants / agriculteurs	Autres	Pas d'exclusivité déclarée
Nombre de contrats	13	3	133	27	1	11	128
Part des assurés collectifs	2 %	0,1 %	45%	15%	0,1%	0,3%	36%
Score moyen	0,85	0,70	0,84	0,80	0,75	0,75	0,80
Prime de référence moyenne	70 €	97 €	66 €	48 €	50 €	113 €	64 €

Lecture > 133 contrats de l'enquête OC 2016 sont déclarés comme étant destinés aux salariés du secteur privé. Ces contrats ont été attribués à 45 % des bénéficiaires de contrats d'entreprise de SRCV 2017. Le score moyen pour ces contrats s'élève à 84 % et la prime de référence moyenne est de 66 euros mensuels.

Champ > Contrats collectifs non surcomplémentaires de l'enquête OC.

Source > Enquête OC 2016. Traitements DREES.

Étant donné leur niveau de garanties et leur tarification, les contrats n'ayant pas déclaré d'exclusivité spécifique semblent être des contrats d'entreprise classiques ayant jugé que la question ne les concernait pas. Les catégories « retraités », « fonctionnaires », « indépendants » et « autres » sont considérées comme ne relevant pas de contrats d'entreprise et sont donc classées comme individuelles.

Imputation des contrats

L'imputation de contrats de complémentaire santé aux individus de SRCV 2017 se fait sur la base des éléments permettant de distinguer entre eux les contrats dans l'enquête OC 2016 :

- type individuel ou d'entreprise ;
- score de qualité de contrat estimé sur la base des garanties offertes, et classé en trois grandes catégories ;
- nombre d'assurés par classe d'âge ;
- exclusivité pour une certaine catégorie de population : retraités, étudiants, salariés du privé ou fonctionnaires ;
- nombre de bénéficiaires par assuré.

Les mêmes strates d'imputation doivent être créées pour les individus de SRCV, c'est-à-dire qu'il faut les catégoriser par classe d'âge, par nature du contrat, par qualité de contrat estimée grâce à la prime déclarée et par types de population susceptibles de donner lieu à une exclusivité. Cela est bien sûr immédiat pour la classe d'âge et le type de contrat.

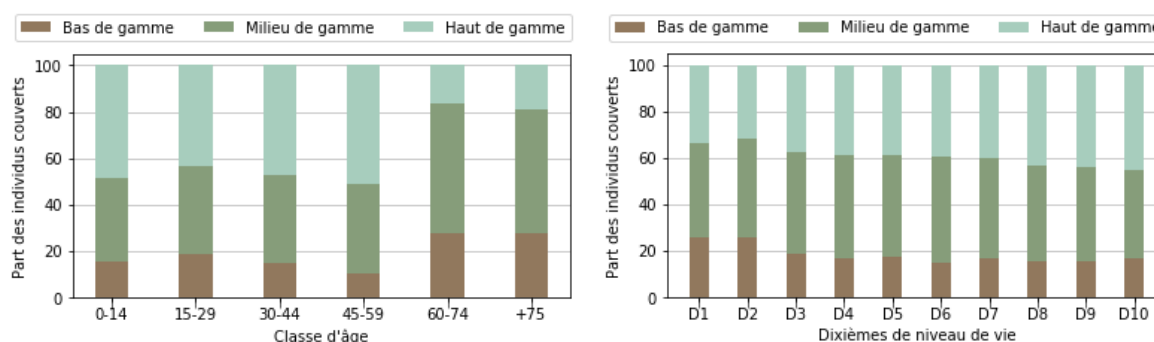
Concernant le score de qualité, la proportion de chaque catégorie de score dans la classe d'âge et de type de contrat peut être connue grâce à l'enquête OC. L'attribution des contrats s'appuie sur l'hypothèse que, au sein d'une même strate, les différences de qualité de contrat sont liées aux différences de cotisations versées. Les cotisations déclarées par les ouvriers droit dans SRCV sont donc individualisées en prenant en compte le nombre et le type de bénéficiaires, afin d'obtenir le meilleur estimateur possible de la qualité du contrat. Puis les individus d'une même classe d'âge et de type de contrat sont classés par ordre croissant de cotisations individualisées, et les scores de qualité sont attribués conformément aux proportions dans la strate. Ceci est une nouveauté importante par rapport aux précédentes versions du modèle, permettant notamment d'améliorer la corrélation entre qualité du contrat, niveau de vie et âge. Le graphique 2 montre la répartition des souscripteurs entre trois gammes de qualité de contrat, définies grâce à la méthode de classification des contrats de la DREES, et selon leur classe d'âge et leur niveau de vie.

Concernant les exclusivités, ces contrats ne doivent être attribués qu'à des individus de la catégorie concernée, mais tout en conservant les proportions observées dans l'enquête OC. Un calcul est fait pour savoir quelle part des retraités (par exemple) est dans des contrats exclusivement réservés aux retraités : pour chaque strate (type de contrat x qualité) la proportion de retraités dans SRCV est mise en regard de la proportion d'assurés de cette strate ayant un contrat réservé aux retraités dans l'enquête OC. Si deux tiers des individus de la strate sont des retraités et qu'un tiers ont un contrat réservé aux retraités, cela signifie que la moitié des retraités dans cette strate sont dans un contrat qui leur est réservé. Cette probabilité est alors respectée lors de l'attribution des contrats de complémentaire santé : la moitié des retraités de la strate vont recevoir un contrat réservé aux retraités, tandis que l'autre moitié sera affectée à des contrats ouverts à tous.

Les contrats qui sont financièrement les plus avantageux pour les familles comptent en moyenne davantage de bénéficiaires par assuré, et inversement. Ne pas prendre en compte cet effet conduirait à surestimer les primes en attribuant à de nombreuses familles des contrats peu avantageux pour elles. Pour chaque contrat de l'enquête OC la proportion d'assurés sans ayant droit peut être estimée grâce au nombre moyen de bénéficiaires par assuré. Cette proportion est ensuite utilisée pour moduler les probabilités d'attribution des contrats selon si l'assuré a ou non des ayants droit.

Chaque ouvrant droit ayant été placé dans la strate qui lui correspond, l'imputation à celui-ci et à ses ayants droit d'un contrat donné parmi ceux de la strate se fait au tirage au sort, en respectant le poids des contrats en termes de nombre d'assurés au sein de la classe d'âge et en termes d'existence ou non d'ayant droit.

Graphique 2 • Répartition des qualités de contrat selon les revenus et l'âge de l'assuré, après imputation



Lecture > 56 % des personnes couvertes par un contrat de complémentaire santé privé et âgées d'entre 61 et 75 ans ont un contrat dont le niveau de couverture est classifié comme « Milieu de gamme ». Ce chiffre est de 44 % pour les personnes couvertes du troisième dixième de niveau de vie.

Champ > Ensemble des personnes assurées par un contrat de complémentaire santé privé en France métropolitaine.

Sources > SRCV 2017 ; enquête OC 2016. Traitements DREES.

■ CALCUL DES COTISATIONS DE COMPLÉMENTAIRE SANTÉ

À présent chaque ouvrant droit de SRCV et ses ayants droit se sont vu attribuer un contrat issu de l'enquête OC. Les primes déclarées dans SRCV ont permis d'estimer la qualité du contrat, permettant ainsi une imputation plus fine. Néanmoins, rien ne garantit qu'elles soient suffisamment fiables pour être retenues telles quelles dans le modèle. Les informations de tarification disponibles dans l'enquête OC sont donc mobilisées pour simuler la prime associée au contrat. Le choix de s'appuyer sur ces informations semble validé par les très bonnes propriétés des distributions de prime obtenues (voir *infra*).

Retraitements des informations de cotisation dans l'enquête OC

Les données de l'enquête OC utilisées sont relativement brutes. Étant donné la complexité des différentes modalités de calcul des cotisations, un nombre important de retraitements préalables sont nécessaires.

- Pour 37 contrats, aucune information de cotisation n'est disponible, ni pour l'individu de référence, ni pour aucun autre type d'assuré. Ces contrats ne sont donc attribués à personne.
- Concernant les contrats individuels, il se peut que les cotisations à certains âges ne soient pas déclarées car le contrat est en théorie interdit aux personnes de cet âge. C'est notamment fréquemment le cas aux âges les plus élevés, ou inversement pour la cotisation à 20 ans. Étant donné que les classes d'âge utilisées lors de l'attribution des contrats sont de grande dimension, il ne peut pas être évité que des contrats de ce type soient attribués à des individus auxquels ils ne sont pas destinés. Il est donc nécessaire de simuler une prime à tout âge. Cela est fait en s'appuyant sur la cotisation la plus proche déclarée, par exemple celle à 75 ans pour remplir celle à 85 ans, à laquelle est appliqué le ratio moyen observé sur les contrats individuels pour lesquels les deux variables sont remplies. Néanmoins, étant donné que l'attribution des contrats se fait en tenant compte du nombre d'assurés du contrat dans chaque classe d'âge, lequel doit par définition être très faible si le contrat n'est pas accessible à une classe d'âge spécifique, ce cas de figure est peu fréquent.
- Quelques contrats d'entreprise déclarant moduler leur tarification selon le revenu n'ont pas renseigné la cotisation pour l'individu de référence, tout en ayant renseigné une cotisation pour la tranche de revenus la plus élevée. Nous supposons dans ce cas que l'absence de réponse est liée à la non-pertinence de la question pour ce contrat, c'est-à-dire au fait que le contrat serait gratuit pour la tranche de revenus la plus basse. On renseigne alors une cotisation nulle pour l'individu de référence. Ce cas de figure ne concerne que trois contrats de complémentaire santé issus de l'enquête OC, ce qui est très anecdotique et n'est donc pas de nature à modifier significativement les résultats issus du modèle.

Simulation fine des cotisations

La cotisation de complémentaire santé d'un ouvrant droit est obtenue en modulant la prime de référence du contrat en fonction des caractéristiques individuelles. Ces caractéristiques à prendre en compte sont :

- l'âge de l'ouvrant droit ;
- la localisation de l'ouvrant droit ;
- les revenus de l'ouvrant droit ;
- le nombre de bénéficiaires (adultes et enfants).

Âge • La prime de référence est celle d'un individu de 40 ans. Les OC ont cependant déclaré les cotisations à plusieurs âges : 20, 60, 65, 75, 85. Entre deux âges déclarés, l'hypothèse est faite que les primes évoluent de façon linéaire. En dessous de 20 ans et après 85 ans il est supposé que les primes n'évoluent pas. Ces hypothèses semblent conformes à ce qui est pratiqué par les OC. La prime modulée par l'âge obtenue fait office de prime de base, à laquelle les autres modulations sont appliquées sous forme de coefficients multiplicateurs.

Localisation • Certains OC modulent leurs primes en fonction de la région de résidence. Pour ces OC, la prime de référence est celle pour un individu situé dans la région la plus chère. Cependant, les OC concernés ont également déclaré ce que serait la cotisation de l'individu de référence s'il résidait dans la zone géographique la moins chère.

La majorité des OC déclarent ne pas moduler les primes selon la localisation, mais cela peut induire en erreur car de nombreux petits OC n'exercent en fait que dans une région précise. Ainsi, un OC n'exerçant qu'en Île-de-France pourrait avoir des cotisations bien supérieures à celles d'un OC n'exerçant qu'en Alsace, tandis que les deux OC déclareraient ne pas moduler leurs primes selon le lieu de résidence. Or la méthode d'attribution des contrats ne permet pas d'éviter qu'un contrat francilien ne soit attribué à un ouvrant droit Alsacien. Pour éviter d'introduire un biais, nous faisons l'hypothèse que les primes dépendent

du reste à charge (RAC) AMO moyen dans la région et que cette logique est suivie par tous les organismes, même s'ils ont déclaré ne pas moduler leurs cotisations selon le lieu de résidence. Ainsi, la prime est modulée pour tous les ouvrants droit selon leur lieu de résidence, en l'occurrence selon la zone d'études et d'aménagement du territoire (ZEAT) où ils habitent³. Les données utilisées pour la modulation géographique sont les RAC AMO moyens pour la tranche d'âge 40-50 ans, issus du SNDS 2016 : les régions où les restes à charge sont les plus élevés se voient attribuer des primes de complémentaire santé plus importantes et inversement.

Revenus • Certains OC modulent leurs primes selon les revenus bruts de l'ouvrant droit, notamment pour les contrats d'entreprise. La prime de référence correspond alors à un salarié à temps plein de la tranche de revenus la plus faible (au smic donc). Pour ces contrats, il est également indiqué la prime pour une personne rémunérée 6 436 euros brut par mois⁴ (ou, à défaut, pour la tranche de revenus la plus élevée). La hausse de la prime entre ces deux extrêmes est considérée linéaire. En dessous du smic et au-delà de 6 436 euros brut par mois, il est considéré que la prime n'évolue plus.

Bénéficiaires • La plupart des contrats comportent un surcoût en cas d'extension du contrat à des ayants droit. Dans ce cas, il est indiqué quelle serait la prime de l'individu de référence s'il couvrait son conjoint, son conjoint et un enfant, et deux enfants, et trois enfants. Il peut être déduit de ces différents éléments quel est le surcoût individuel du conjoint et de chaque enfant. Chaque situation individuelle, même celle des familles monoparentales, peut ainsi être simulée en cumulant les surcoûts à la prime de l'ouvrant droit.

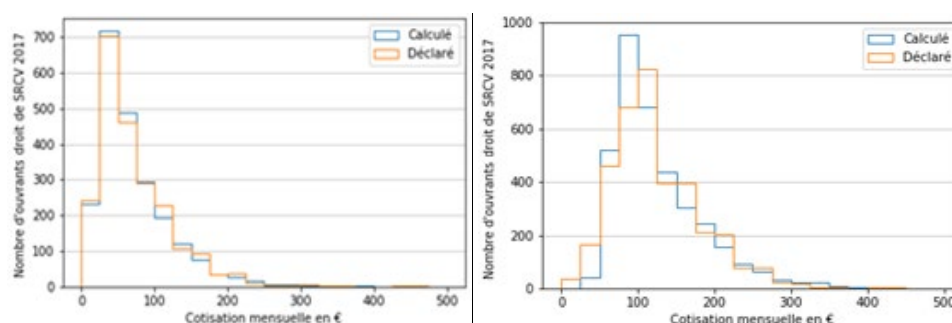
Ici, l'âge pris en compte pour la modulation du surplus de prime est celui des adultes bénéficiaires, et non pas l'âge de l'ouvrant droit lui-même, généralement plus élevé, sans quoi les cotisations seraient surestimées. En outre, contrairement à la localisation et au revenu, le correctif bénéficiaire est sommé à la prime calculée, et non pas appliqué à elle sous forme de multiplicateur, afin de ne pas dépendre de l'âge de l'assuré. La localisation et les revenus sont toutefois pris en compte sous forme de multiplicateurs appliqués au surplus de prime.

Vérifications

La distribution des primes calculées à partir des contrats imputés peut être comparée à celle des primes déclarées dans SRCV. Dans l'ensemble, les deux distributions sont très proches, et ce même en restreignant à des sous-populations précises. Cette cohérence valide à la fois la fiabilité des primes déclarées dans SRCV et la méthode employée pour simuler des primes à partir de l'enquête OC. Nous faisons le choix de conserver la prime simulée à partir de l'enquête OC afin qu'il y ait cohérence entre la prime payée par un ménage et le niveau de garanties dont il bénéficie, lui aussi issu de l'enquête OC.

Le graphique 3 compare les distributions des primes déclarées dans SRCV avec celles recalculées finement. Le champ est restreint aux contrats individuels. À gauche, pour les ouvrants droit de moins de 50 ans, à droite pour les ouvrants droit de plus de 65 ans. Pour ces deux populations, on constate que les distributions tout comme les moyennes sont très semblables. Il en va de même pour les autres sous-populations analysées : assurés n'ayant aucun ayant droit ; assurés ayant au moins un ayant droit ; personnes des trois premiers dixièmes de niveau de vie ou des trois derniers dixièmes.

Graphique 3 • Distributions comparées des primes déclarées dans SRCV et recalculées pour les détenteurs de contrats individuels de moins de 50 ans (gauche) et de plus de 65 ans (droite)



Lecture > Parmi les souscripteurs de contrats individuels âgés de moins de 50 ans, 467 personnes enquêtées dans SRCV 2017 ont déclaré une cotisation de complémentaire santé située entre 50 et 75 euros par mois. Ils seraient 494 à se situer dans cette tranche de cotisation d'après les calculs fins effectués dans le cadre du modèle.

Champ > Ensemble des ouvrants droit de contrats individuels de moins de 50 ans (gauche) et de plus de 65 ans (droite).

Sources > SRCV 2017 ; enquête OC 2016. Traitements DREES.

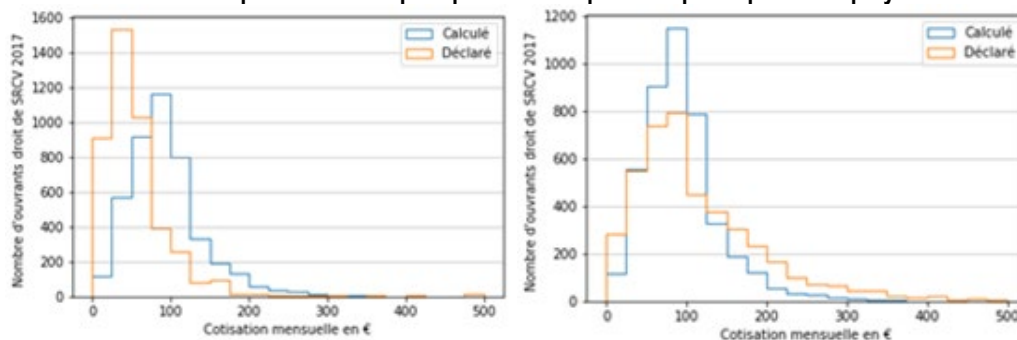
Concernant les contrats d'entreprise, si l'on trace les distributions telles que déclarées dans SRCV, un gros décalage existe avec les primes calculées (graphique 4 gauche). Les enquêtés ont déclaré combien ils paient pour leur complémentaire santé

³ Les ZEAT sont des regroupements de départements tels que « Bassin parisien », « Sud-Ouest » ou « Méditerranée ».

⁴ Soit deux fois le plafond mensuel de la Sécurité sociale de 2016.

et non pas quel est le coût réel. Pour pouvoir effectuer la comparaison, il est impératif d'ajouter aux primes déclarées la part prise en charge par l'employeur. Les résultats de l'enquête Protection sociale complémentaire d'entreprise sont mobilisés pour affecter une part moyenne de prise en charge par l'employeur dépendante du secteur d'activité déclaré dans SRCV (Barlet, *et al.*, 2019 ; Perronin, 2019). À la suite de cette correction, les deux distributions sont nettement plus cohérentes entre elles (graphique 4 droite). Néanmoins, des divergences existent, avec notamment un certain nombre d'ouvrants droit dont la cotisation déclarée corrigée dépasse la cotisation simulée. La part prise en charge par l'employeur au sein même d'un secteur d'activité peut fortement varier d'un salarié à l'autre, de sorte que la correction a pu attribuer des participations employeurs surévaluées à une partie de la population.

Graphique 4 • Distributions comparées des primes déclarées dans SRCV et recalculées pour les détenteurs de contrats d'entreprise avant et après prise en compte de la participation employeur



Note > À gauche, les primes déclarées sont directement issues de SRCV, à droite les primes déclarées ont été corrigées de la participation employeur.

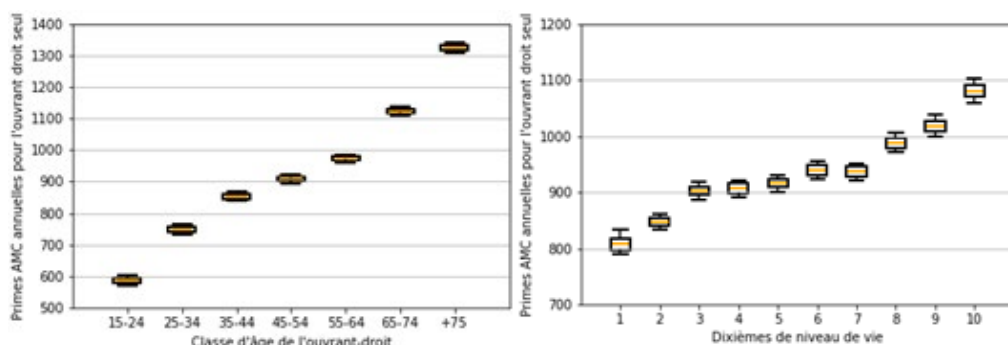
Lecture > Parmi les ouvriers droit de contrat d'entreprise, 400 enquêtés de SRCV 2017 ont déclaré une cotisation de complémentaire santé située entre 75 et 100 euros par mois (graphique de gauche). Les simulations effectuées dans le cadre du modèle, et qui incluent la part financée par l'employeur, comptabilisent 1183 ouvriers droit dans cette tranche de cotisations.

Champ > Ensemble des ouvriers droit de contrats d'entreprise.

Sources > SRCV 2017 ; enquête OC 2016. Traitements DREES.

La méthode employée pour l'imputation des contrats comporte une part d'aléas : pour chaque ouvrant droit, un contrat de sa strate est tiré au sort. Les garanties et le montant des cotisations retenus pour ce dernier sont alors ceux du contrat tiré. Il est légitime de s'inquiéter qu'un autre tirage au sort conduise à avoir des résultats différents. Afin d'étudier la variabilité des résultats, l'opération d'attribution des contrats et de simulation des cotisations est effectuée 100 fois. Ceci permet d'estimer 100 fois les moyennes de cotisation par dixième de niveau de vie et par classe d'âge. Les boîtes à moustaches correspondantes sont données en graphique 5.

Graphique 5 • Boîtes à moustaches de répartition des 100 itérations, moyennes de cotisation par classe d'âge (gauche) et par dixième de niveau de vie (droite)



Lecture > Sur 100 itérations de l'algorithme d'attribution des contrats de complémentaire santé de l'enquête OC 2016 aux individus de SRCV 2017, concernant la cotisation moyenne des individus du 6^e dixième de niveau de vie, le premier décile se situe à 923 euros par an, le premier quartile à 930 euros par an, la médiane à 939 euros, le troisième quartile à 947 euros et le neuvième décile à 954 euros.

Champ > Ensemble des personnes en ménage ordinaire de France métropolitaine.

Sources > SRCV 2017 ; enquête OC 2016. Traitements DREES.

Concernant les moyennes par classe d'âge, on constate que la variabilité est très faible et qu'il n'y a pas de risque d'obtenir des moyennes erronées. Pour ce qui est de la moyenne par dixième de niveau de vie, la variabilité est également assez faible, l'écart interquartiles étant de l'ordre de quelques dizaines d'euros par an. Néanmoins, quelques valeurs peuvent s'éloigner fortement au-delà des premier et neuvième déciles affichés dans le graphique, ce qui fait courir un léger risque d'obtenir une moyenne erronée jusqu'à hauteur d'une centaine d'euros.

ESTIMATION DES REMBOURSEMENTS DE COMPLÉMENTAIRE SANTÉ

À présent, les individus de l'enquête SRCV se sont vu imputer des dépenses de santé issues du SNDS, ainsi que des contrats de complémentaires santé incluant des garanties détaillées. Le niveau de finesse de ces deux ensembles de données imputées permet une microsimulation détaillée des remboursements que ces individus perçoivent de la part de leur complémentaire santé.

Modes d'expression des garanties

Pour chaque poste de soins, l'enquête OC propose aux complémentaires santé de détailler les garanties proposées sous plusieurs formats possibles. Le remboursement peut ainsi s'exprimer en pourcentage du ticket modérateur, proportion pouvant excéder 100 % si le contrat rembourse des dépassements d'honoraires. Il peut aussi s'exprimer en pourcentage des frais réels, c'est-à-dire en pourcentage de l'ensemble de la dépense de santé. Il est également fréquent que le remboursement puisse s'exprimer sous forme d'un forfait monétaire par acte, ou d'un forfait annuel. Certains contrats proposent un remboursement forfaitaire pour ce qui dépasse du pourcentage remboursé sur le ticket modérateur ou sur les frais réels. Par exemple, un contrat peut rembourser l'intégralité du ticket modérateur d'une visite chez le spécialiste ainsi que les dépassements d'honoraires à hauteur de 20 euros.

Tableau 3 • Modes d'expression des garanties possibles dans l'enquête OC selon le poste de soins

Poste de soins	RAC opposable si responsable	% du ticket modérateur	% des frais réels	Forfait monétaire acte par acte	Forfait annuel	Plafond éventuel
Hôpital frais de séjour	x					
Honoraires de praticiens hospitaliers		x	x	x		
Généralistes et auxiliaires	x		x	x		
Médecins spécialistes		x	x	x		
Radiologie		x	x			
Médicaments		x				
Analyses biologiques		x	x			
Transports et matériel	x					
Monture et verres		x	x	x		
Lentilles					x	
Soins dentaires conservateurs		x	x	x		
Prothèses dentaires		x	x	x		x
Orthodontie		x	x		x	x
Implantologie / parodontologie			x			
Prothèses auditives			x		x	

Lecture > Les dépenses de santé chez le médecin spécialiste peuvent être remboursées sous la forme soit d'un pourcentage du ticket modérateur, soit d'un pourcentage de l'ensemble de la dépense, soit d'un forfait monétaire pour chaque acte. Le mode de remboursement opéré dans le modèle dépend de ce qui a été déclaré par l'organisme complémentaire dans l'enquête OC pour le contrat imputé.

Source > Enquête OC 2016.

Concernant l'optique, les OC ont la possibilité de rembourser l'ensemble monture + verres avec un unique forfait, lequel diffère selon la nature simple ou complexe des verres achetés. Dans les très rares cas où l'individu aurait consommé un verre simple et un verre complexe sur une même monture, l'ensemble est considéré comme étant de l'optique complexe. Enfin, pour les prothèses dentaires et l'orthodontie, les organismes complémentaires peuvent préciser un plafond annuel de remboursement

ne pouvant pas être dépassé. Le tableau 3 récapitule les différents modes d'expression des garanties possibles dans l'enquête OC pour chaque poste de soins.

Niveau de détail des simulations

Les simulations de remboursement s'efforcent de toujours mobiliser le niveau de détail le plus fin disponible dans l'enquête OC. En particulier, pour les postes où les garanties peuvent être exprimées sous forme de forfait monétaire, les dépenses de santé ont été conservées à un niveau de détail acte par acte lors de l'agrégation des données du SNDS. De même, la simulation respecte le niveau de détail proposé par l'enquête OC sur les garanties appliquées aux consultations de spécialiste, lesquelles sont différenciées selon que la visite respecte ou non le parcours de soins et selon que le spécialiste adhère ou non au contrat d'accès aux soins (CAS).

Pour certains postes de soins, l'enquête OC ne donne aucune précision sur les garanties appliquées : frais de séjour à l'hôpital, médecins généralistes, sages-femmes, auxiliaires, transports, analyses biologiques, matériel médical. Pour ces postes, il est considéré que les contrats responsables remboursent l'intégralité du RAC opposable mais ne remboursent pas les franchises médicales. Les dépassements d'honoraires éventuels sont traités comme s'il s'agissait de consultations de spécialistes. Concernant les contrats non responsables, lesquels couvrent moins de 2 % de la population, l'enquête OC n'indique pas le niveau de remboursement offert pour ces postes. Il est alors supposé que ces contrats ne remboursent pas le RAC opposable.

Les organismes complémentaires ont la possibilité de déclarer des niveaux de garantie différents selon que le soin a été consommé dans ou hors du réseau de professionnels qui leur est propre. Il n'est bien entendu pas possible de savoir si une certaine dépense de santé observée dans les données a été effectuée dans ou en dehors du réseau de soins lié à la complémentaire santé du consommant. L'hypothèse faite est que l'ensemble des dépenses de santé ont lieu au sein du réseau lorsque celui-ci existe. Ceci est en accord avec l'objectif d'étudier le remboursement maximal auquel les individus peuvent prétendre s'ils font les démarches idoines.

Suite aux travaux menés sur la couverture complémentaire santé des individus de SRCV (voir partie II), certains individus du modèle peuvent être couverts par deux contrats différents. Le remboursement perçu sur un acte ou sur un poste agrégé de dépenses est alors simulé pour chacun des deux contrats avant que ne soit retenu le meilleur remboursement.

■ LE FINANCEMENT DE L'ASSURANCE MALADIE OBLIGATOIRE

À ce stade de la construction du modèle il est possible de simuler les dépenses de santé, les remboursements par l'AMO, les cotisations de complémentaire santé, les remboursements versés par les OC et donc ce qui reste à la charge des ménages après intervention des complémentaires. Le modèle Omar est donc achevé. Néanmoins, il ne permet pas à lui seul d'avoir une vision d'ensemble de la redistribution opérée par le système de santé. Pour ce faire, un pseudo-appariement au modèle Ines 2017, géré conjointement par l'Insee, la DREES et la CNAF, est nécessaire afin de connaître la contribution de chaque individu au financement de l'AMO.

Pseudo-appariement à Ines 2017

Le sens de l'imputation

Deux stratégies d'imputation sont possibles : soit transmettre aux individus du modèle Omar des données de contribution au financement de l'AMO issues d'un donneur côté Ines ; ou, inversement, transmettre aux individus du modèle Ines les dépenses de santé et les éléments de complémentaire santé issus du modèle Omar. La première méthode a l'avantage de transférer un plus faible nombre de variables. Néanmoins, le modèle Ines comptant environ cinq fois plus d'individus que le modèle Omar, la seconde méthode apporte plus de puissance statistique en permettant de mobiliser l'ensemble des données d'Ines.

En outre, les méthodes d'imputation reposant nécessairement sur une part de hasard, le fait que chaque individu d'Omar soit apparié en moyenne à cinq individus d'Ines permet de réduire l'influence de l'aléa par rapport à une solution inverse où seul un individu d'Ines sur cinq aurait été apparié à un individu d'Omar. C'est donc la seconde option qui est retenue. Le modèle Omar peut ainsi être considéré comme un module complémentaire du modèle Ines, d'où l'appellation « Ines-Omar ».

Étant donné que plusieurs étapes de construction du modèle Omar comportent également de l'aléa (attribution d'un contrat de complémentaire santé et pseudo-appariement EHS-SRCV notamment), cette logique de « *bootstrap* » est poussée un cran plus loin. Au point de départ de la construction, la base SRCV est dupliquée en un total de cinq exemplaires pour obtenir une base d'une taille comparable au modèle Ines. Les cinq individus identiques de SRCV vont alors connaître des parcours divergents : pseudo-appariement avec des individus d'EHS différents, attribution de contrats de complémentaire santé différents, recours ou non à l'ACS lorsqu'il y a éligibilité, etc. Au moment du pseudo-appariement avec Ines, ces cinq individus représentent donc autant de réalisations divergentes de quelques-uns des multiples parcours possibles, réduisant par conséquent l'influence du hasard dans la construction du modèle d'ensemble final.

Le pseudo-appariement par « *hot-deck* »

L'imputation de données d'Omar aux individus d'Ines étant actée, il reste à choisir la méthode d'imputation. Le nombre de variables à transférer est élevé et ces variables interagissent fortement entre elles : les dépenses de santé ont une structure très différente en cas d'ALD ou de recours à la CMU-C par exemple ; la cotisation de complémentaire santé et les remboursements perçus sont directement corrélés ; la base de remboursement doit correspondre à la somme du ticket modérateur et du remboursement AMO ; le reste à charge AMC doit correspondre au reste à charge opposable moins les remboursements AMC, etc. Des méthodes de régression ou d'apprentissage statistique qui chercheraient à imputer dans Ines les variables une par une par simulation perdraient partiellement ces corrélations, lesquelles sont essentielles aux analyses qui seront menées à l'aide du modèle. C'est pourquoi il est choisi de pseudo-apparier des individus et de transmettre telles quelles l'ensemble des données d'un donneur vers un receveur.

Les données à transmettre sont de plusieurs ordres : état de santé, consommations de soins et couverture complémentaire. Tout en étant partiellement corrélés, ces éléments ont une part d'indépendance. Or les méthodes d'apprentissage statistique nécessitent de sélectionner une variable expliquée à l'exclusion des autres. Il est possible d'essayer d'optimiser plusieurs variables simultanément, mais cela rend l'algorithme particulièrement gourmand en données. Or l'échantillon disponible est de dimension plutôt restreinte (24 000 individus dans SRCV) et les variables communes aux deux bases à pseudo-apparier sont certes nombreuses mais très redondantes : la plupart concernent l'estimation très fine des revenus ou du statut d'occupation. C'est pourquoi les méthodes d'apprentissage statistique ne semblent pas appropriées pour ce pseudo-appariement.

A l'inverse, la méthode du « *hot-deck* » ne cherche pas à apprendre mais laisse au statisticien la tâche de choisir les variables explicatives qui lui semblent les plus pertinentes et leur ordonnancement. Cette démarche est plus adaptée à un contexte où de nombreuses variables sont à expliquer à partir de peu de variables explicatives. Dans le *hot-deck*, le donneur est tiré aléatoirement parmi une population de donneurs ayant les mêmes caractéristiques que le receveur. Afin de réduire l'influence

de l'aléa, les strates parmi lesquelles le tirage est effectué doivent avoir une taille minimale. Si, en mobilisant l'ensemble des variables utilisées pour l'imputation, certaines strates ne sont pas de taille suffisante, les individus faisant partie de cette strate seront alors pseudo-appariés lors d'un deuxième tour, en supprimant l'une des variables d'imputation. On procède ensuite à un troisième tour, puis à un quatrième, etc. en supprimant chaque fois une nouvelle variable, et ce jusqu'à ce que tous les individus soient pseudo-appariés.

Il est donc nécessaire de sélectionner les variables retenues pour la stratification ainsi que l'ordre dans lequel elles seront retirées. Nous sélectionnons les variables communes aux deux modèles les plus importantes pour les analyses que nous souhaitons mener : classe d'âge en tranches décennales, sexe, statut d'occupation (salarié du secteur privé, retraité, étudiant, etc...), niveau de vie par vingtièmes et type de couverture complémentaire santé. Cette dernière variable ne permet que de distinguer les personnes à la CMU-C et à l'ACS, la couverture privée étant inconnue dans le modèle Ines. Le recours à la CMU-C et à l'ACS a été attribué aléatoirement aux personnes éligibles, en respectant les proportions à âge et sexe donnés⁵. Vu le grand nombre de données à transférer (statut ALD, dépenses de santé, couverture complémentaire, etc.), il est impensable d'ordonner les variables en fonction de leur significativité vis-à-vis d'une variable précise. En effet, par exemple, les caractéristiques individuelles importantes pour savoir si une personne est en ALD ou non ne sont vraisemblablement pas tout à fait identiques aux caractéristiques importantes pour comprendre la qualité de la complémentaire santé souscrite. Les variables d'imputation sont donc ordonnées selon leur ordre d'importance pour les analyses que l'on désire mener avec le modèle : type de couverture en premier, puis niveau de vie, âge, statut d'occupation et enfin sexe.

La taille minimale d'une strate est de cinquante donneurs potentiels. Comme le montre le tableau 4, avec ce paramétrage le *hot-deck* apparie plus des trois quarts des individus d'Ines en conservant l'ensemble des cinq variables de stratification, et 95 % d'entre eux avec au moins les variables de type de couverture complémentaire, de niveau de vie et d'âge.

Tableau 4 • Déroulement de l'imputation par *hot-deck* des données d'Omar vers Ines

		1 ^{er} tour	2 ^e tour	3 ^e tour	4 ^e tour	5 ^e tour
Nombre et part d'individus d'Ines appariés		88 793	7 667	14 834	5 202	846
		75,7 %	6,5 %	12,6 %	4,4 %	0,7 %
Variables de stratification	Type de couverture	x	x	x	x	x
	Niveau de vie	x	x	x	x	
	Classe d'âge	x	x	x		
	Statut d'occupation	x	x			
	Sexe	x				

Lecture > 7667 individus du modèle Ines ont été pseudo-appariés lors du deuxième tour, soit 6,5 % des individus. Toutes les variables de stratification sauf le sexe ont été mobilisées dans ce deuxième tour.

Champ > Ensemble des individus du modèle Ines 2017.

Sources > Ines 2017 ; Omar 2017. Traitements DREES.

Il est important de préciser que le pseudo-appariement est effectué entre individus et non entre ménages. Il n'y a donc dans Ines-Omar aucune garantie que la structure de couverture complémentaire soit cohérente au sein d'un ménage : un ménage peut par exemple se composer d'une personne considérée ayant droit d'un contrat d'entreprise et d'une autre couverte par un contrat individuel. Ceci n'est pas problématique tant que l'on mène des études sur une sous-population, telle qu'une certaine catégorie de niveau de vie, une certaine classe d'âge ou un certain statut d'emploi. En revanche, si l'on souhaite étudier plus précisément la situation de couverture à l'intérieur des ménages il est important de revenir au modèle Omar non pseudo-apparié à Ines.

Calages macro-économiques

Le modèle Ines-Omar nécessite un calage des montants de dépense de santé pour trois raisons. En premier lieu, les données de santé du modèle sont datées de 2014, tandis que le reste du modèle est conçu pour correspondre à la situation de 2017. Étant donné la croissance rapide des dépenses de santé, il est impératif de mettre à jour les montants à la date de 2017 pour pouvoir se prononcer sur leur niveau. En deuxième lieu, les données de santé disponibles ne concernent qu'un peu plus de 16 000 individus. Cela est insuffisant pour garantir que les niveaux de dépense agrégés estimés soient corrects en raison de

⁵ Ces proportions par âge et par sexe sont issues de SRCV 2017 pour la CMU-C, et du SNDS 2017 pour l'ACS.

la variance associée à un échantillon de taille restreinte. Enfin, en troisième lieu, les personnes hospitalisées au moment de l'enquête, donc ayant vraisemblablement de très fortes dépenses de santé, ne sont pas présentes dans les données.

Pour chaque poste de soins, le montant agrégé de dépense est calé sur une version retraitée des comptes de la santé 2017 : seules sont conservées les dépenses remboursables, individualisables et incluses dans le champ du modèle (Gonzalez, *et al.*, 2018). Le tableau 5 compare pour chaque poste de soins les montants de dépense agrégés avant et après calage, ainsi que ceux issus des comptes de la santé non retraités. Les cotisations de complémentaire santé sont également calées à la baisse en appliquant le même différentiel que celui retenu sur la dépense totale. Ceci permet de prendre en compte l'absence dans le modèle des dépenses non remboursables, non individualisables, ainsi que de certains postes de soins tels que les cures thermales ou la chambre particulière à l'hôpital.

Le calage ne concerne que les niveaux des montants. La répartition de la dépense entre les individus, ainsi que la décomposition de la dépense entre remboursement AMO, remboursement AMC, liberté tarifaire, RAC opposable, etc. sont ceux estimés dans le modèle sans calage. Le tableau 6 présente les contributions de chacun des trois financeurs respectivement dans le modèle Ines-Omar 2017 après calage et dans les comptes de la santé 2017 non retraités. Les écarts observés sont en bonne partie dus à des différences de champ. En effet, les dépenses de santé qui ne figurent pas dans le modèle sont généralement moins bien remboursées par l'AMO, donc reposent plus fortement sur les ménages, ce qui aboutit à une répartition moyenne du financement différente.

Tableau 5 • Dépense agrégée de santé avant et après calage du modèle Ines-Omar 2017, et dans les comptes de la santé 2017

Poste de soins	Avant calage	Après calage	Comptes 2017
Hôpital public	45 Mds €	59,3 Mds €	71,5 Mds €
Hôpital privé	11,7 Mds €	15,1 Mds €	21,4 Mds €
Médecins et sages-femmes	17,1 Mds €	20,1 Mds €	21,2 Mds €
Auxiliaires médicaux	14,1 Mds €	14,5 Mds €	15,8 Mds €
Soins dentaires	8,9 Mds €	9,9 Mds €	11,3 Mds €
Laboratoires d'analyses	4,1 Mds €	4,5 Mds €	4,5 Mds €
Transports sanitaires	4,0 Mds €	4,9 Mds €	5,0 Mds €
Médicaments	29,8 Mds €	29,1 Mds €	32,6 Mds €
Autres biens médicaux	13,3 Mds €	15 Mds €	15,5 Mds €
Total	148 Mds €	172,4 Mds €	198,8 Mds €

Lecture > Avant calage macro-économique, le modèle estime le total de dépenses de santé en soins dentaires à 8,9 milliards d'euros sur l'année 2017. Après calage macro-économique sur les données des comptes de la santé adaptées au champ de l'individualisable et du remboursable, le total de dépenses en soins dentaires est de 9,9 milliards d'euros en 2017. Les comptes de la santé non restreints à ce champ comptabilisent 11,3 milliards d'euros de soins dentaires en 2017.

Champ > Ensemble des dépenses de santé individualisables et remboursables pour les colonnes « Avant calage » et « Après calage » ; ensemble des dépenses de santé pour les comptes de la santé.

Sources > Ines-Omar 2017 ; Comptes de la santé 2018. Traitements DREES.

Les prestations des OC prises en compte dans le modèle ne sont que celles qui concernent les soins remboursables par l'AMO et excluent donc une partie des prestations telles que les chambres particulières à l'hôpital, l'homéopathie, la prévention, etc. Il est nécessaire de mettre en cohérence le champ des cotisations d'OC, qui correspondent implicitement à l'ensemble du champ des prestations des OC. Pour cela, le total de cotisations est calé proportionnellement au total de prestations prises en compte dans le modèle. L'écart entre prestations et cotisations correspond aux taux de frais de gestion et aux marges des OC tels qu'estimés par le rapport sur la situation financière des organismes complémentaires assurant une couverture santé portant sur l'année 2017 (Adjerad, 2018). Ces taux de frais et de marges sont différenciés pour les contrats individuels (27 %) et d'entreprise (15 %). Le total de cotisations hors taxes passe ainsi de 38,6 milliards d'euros avant calage à 30,9 milliards d'euros après calage. Ces chiffres apparaissent comme légèrement supérieurs aux estimations du rapport OC, en raison principalement d'une surreprésentation dans le modèle Ines-Omar des contrats individuels, plus onéreux que les contrats collectifs. Cette surreprésentation est elle-même avant tout expliquée par une différence de champ, le modèle opérant une distinction entre les contrats individuels et les contrats d'entreprise, là où le rapport OC distingue l'individuel du collectif, ce dernier étant plus large que les seuls contrats d'entreprise car certains contrats collectifs sont souscrits par des regroupements de retraités ou d'indépendants.

Tableau 6 • Dépense agrégée de santé, contributions des trois financeurs et total de primes AMC dans le modèle Ines-Omar 2017, dans les comptes de la santé 2017 et dans le rapport OC 2018 (cotisations AMC)

	Ines-Omar 2017	Comptes 2017
Dépenses totales	179 Mds €	199 Mds €
Part de remb. AMO	81,4 %	79,3 %
Part de remb. AMC	13,3 %	13,2 %
RAC AMC	5,4 %	7,5 %
Cotisations AMC hors taxes	30,9 Mds €	36,9 Mds €

Lecture > Le modèle Ines-Omar 2017 estime que 80,7 % des dépenses de santé individualisables et remboursables sont remboursées par l'assurance maladie obligatoire. Les comptes de la santé estiment que 79,3 % de toutes les dépenses de santé sont remboursées par l'AMO en 2017.

Champ > Ensemble des dépenses de santé individualisables et remboursables pour Ines-Omar 2017 ; ensemble des dépenses de santé dans les comptes.

Sources > Ines-Omar 2017 ; comptes de la santé 2018, rapport OC 2018.

Distribuer le financement de l'AMO

Le modèle Ines apporte toutes les informations nécessaires au calcul de la contribution des ménages au financement de l'AMO. Il contient en effet, pour chaque ménage, les montants de prélèvement suivants : cotisations sociales maladies (salariales, patronales, des indépendants et des inactifs), cotisation sociale généralisée (CSG) et contribution pour le remboursement de la dette sociale (CRDS). En outre, le module complémentaire sur la taxation indirecte, développé en 2016, fournit pour chaque ménage une estimation des taxes payées au titre de la taxe sur la valeur ajoutée (TVA), des accises sur le tabac et des taxes sur l'alcool (André, Biotteau et Duval, 2016). L'ajout de ce module complémentaire est une nouveauté par rapport aux versions précédentes du modèle, d'autant plus importante que le profil de ces taxes indirectes est nettement moins progressif que le reste des sources de financement. Ces données issues d'Ines s'ajoutent aux sources de financement qui peuvent être estimées à partir du seul modèle Omar : la taxe de solidarité additionnelle (TSA), qui s'applique proportionnellement aux primes de complémentaire santé, et la taxe sur les médicaments, qui pèse *in fine* sur les RAC AMO des ménages en pharmacie.

Grâce à ces données, il devient possible de distribuer l'ensemble du financement de l'AMO sur les ménages. L'objectif est de savoir comment sont financés les remboursements AMO pris en compte dans le champ du modèle, soit un total de 145,8 milliards d'euros, afin que celui-ci soit équilibré entre ses prestations et ses financements. Par conséquent le financement des dépenses de santé non prises en compte dans le modèle, notamment les dépenses non individualisables telles que les frais généraux de fonctionnement des hôpitaux ou les coûts de gestion de la CNAM, doit être écarté. Il n'existe pas d'affectation spécifique de certaines sources de financement en direction de certains types précis de prestations. C'est pourquoi l'ensemble des sources de financement de l'AMO doivent être prises en compte, y compris la génération de dettes, et réparties proportionnellement à la part qu'elles représentent dans le budget de la Sécurité sociale en 2017. Les comptes de recettes nettes 2017 de la Sécurité sociale⁶ sont mobilisés pour estimer la part du financement issue de chaque source. Ces différentes sources, ainsi que leurs équivalents mobilisés dans le modèle, sont présentés dans le tableau 7. La part calculée dans le tableau exclut les quelques recettes que le modèle ne permet pas de distribuer, lesquelles ne comptent que pour 4 % du financement total de l'Assurance maladie.

Quelques précisions sont nécessaires pour bien comprendre le tableau :

- dans les comptes de la Sécurité sociale, l'écart entre les produits et les charges correspond à de l'endettement, lequel doit alors être pris en compte en tant que source de financement. Le remboursement de la dette s'effectuant essentiellement par la CRDS et la CSG, celles-ci sont donc des sources de recettes à hauteur de 4,9 milliards d'euros en 2017 ;
- en 2017, la branche Accidents du travail et maladies professionnelles (AT-MP) a contribué au financement de la branche maladie à hauteur d'un milliard d'euros. Cette contribution est financée conformément au financement du risque AT-MP, c'est-à-dire presque exclusivement par des cotisations salariales⁷ ;
- la Caisse nationale de solidarité pour l'autonomie (CNSA) finance certaines dépenses de santé relevant de la branche maladie tout en recevant des fonds de celle-ci. Au global, il s'agit d'une contribution nette de la CNSA vers la branche maladie, c'est-à-dire d'une source de financement. Hormis les versements par la branche maladie, le financement de la

⁶ Plus précisément, le tableau 2 en partie 4.2 du rapport de Juin 2018 de la Commission des comptes de la Sécurité sociale.

⁷ Voir le tableau 1 de la partie 4.3 du rapport de juin 2018.

CNSA repose essentiellement sur les contributions de solidarité pour l'autonomie (CSA à 76 % et CASA à 24 %) ⁸. Ces contributions ne sont pas connues dans le modèle Ines, mais la CSA est proche des cotisations sociales maladie patronales tandis que la contribution additionnelle CASA s'applique uniformément aux pensions de retraite et d'invalidité. C'est donc la distribution de ces cotisations patronales et de ces pensions que nous prenons comme équivalents à la CSA et à la CASA dans le modèle.

Tableau 7 • Sources de financement de l'assurance maladie obligatoire prises en compte et leur équivalent mobilisé dans le modèle

Source de financement	Équivalent dans Ines-Omar	Montant comptes de la Sécurité sociale	Montant Ines-Omar	Part
Cotis. sociales des salariés	Cotis. sociales maladie des salariés	84,4 Mds €	62,1 Mds €	42,6 %
Cotis. sociales des non-salariés	Cotis. sociales maladie des non-salariés	3,8 Mds €	2,8 Mds €	1,9 %
Cotis. sociales des inactifs	Cotis. sociales maladie des inactifs	0,9 Md €	0,6 Md €	0,4 %
Cotisation sociale généralisée affectée à la maladie	CSG	71,2 Mds €	52,3 Mds €	35,9 %
Taxe sur les salaires	Cotis. sociales maladie patronales	3,1 Mds €	2,3 Mds €	1,6 %
Taxe de solidarité additionnelle	Primes OC	2,2 Mds €	1,6 Mds €	1,1 %
Taxe sur les médicaments	RAC AMO en pharmacie	1,0 Md €	0,7 Md €	0,5 %
TVA + produits alimentaires	Module de taxation indirecte	11,2 Mds €	8,3 Mds €	5,7 %
Accises sur le tabac		11,8 Mds €	8,7 Mds €	6,0 %
Taxe sur l'alcool		1,6 Mds €	1,2 Mds €	0,8 %
Transfert de la branche AT-MP	Cotis. sociales maladie des salariés	1,0 Md €	0,7 Md €	0,5 %
Contribution nette à la CNSA	Cotis. sociales maladie patronales (76 %), retraites et invalidité (24 %)	1,3 Mds €	0,9 Md €	0,6 %
Endettement	CRDS (47 %) et CSG (53 %) ⁹	4,9 Mds €	3,6 Mds €	2,5 %
Total		198,4 Mds €	145,8 Mds €	100 %

Lecture > La taxe sur les salaires est approximée dans le modèle comme étant distribuée au sein de la population de la même façon que les cotisations sociales maladie patronales. Elle rapporte un total de 3,1 milliards d'euros en 2017, soit 1,6 % du total des contributions prises en compte dans le modèle, de sorte qu'elle finance les prestations AMO prises en compte dans le champ du modèle à hauteur de 2,3 milliards d'euros.

Champ > Ensemble des contributions au financement de l'assurance maladie obligatoire.

Sources > Ines-Omar 2017 ; comptes de la Sécurité sociale 2018. Traitements DREES.

⁸ Voir le budget exécuté 2017 de la CNSA.

⁹ D'après le rapport annuel 2017 de la Caisse d'amortissement de la dette sociale (CADES), ses ressources ont été de 7,2 milliards d'euros de CRDS, 7,9 milliards d'euros de CSG et un versement de 2,1 milliards d'euros par le Fonds de réserve pour les retraites (FRR). Étant donné que les ressources du FRR proviennent d'investissements économiques fructueux, ils ne peuvent pas être distribués entre les ménages. Par ailleurs, en 2017, cette contribution compense presque parfaitement les intérêts versés par la CADES à ses investisseurs. Il est donc considéré que l'endettement est entièrement financé par la CRDS et la CSG.

■ POUR EN SAVOIR PLUS

Adjerad, R. (2018, décembre). [La situation financière des organismes complémentaires assurant une couverture santé. Rapport 2018](#). Paris, France : DREES.

Adjerad, R., Courtejoie, N. (2021), [Hospitalisation : des restes à charge après AMO plus élevés en soins de suite et de réadaptation et en psychiatrie qu'en médecine, chirurgie, obstétrique et odontologie](#). DREES, *Études et Résultats*, 1192.

André, M., Biotteau, A.-L., Duval, J. (2016), [Module de taxation indirecte du modèle Ines – Hypothèses, principes et aspects pratiques](#). DREES, *Documents de travail*, 60.

Barlet, M., Gaini, M., Gonzalez, L., Legal, R. (dir.) (2019), [La complémentaire santé – Acteurs, bénéficiaires, garanties – Édition 2019](#). Paris, France : DREES, coll. Panoramas de la DREES-Santé.

Fouquet, M. (2020), [Une hausse modérée de la couverture de la population après la généralisation de la complémentaire santé d'entreprise](#). DREES, *Études et Résultats*, 1166.

Fredon, S., Sicsic, M. (2020), [Ines, le modèle qui simule l'impact des politiques sociales et fiscales](#). Insee, *Courrier des statistiques*, 4.

Gonzalez, L., Roussel, R., Héam, J.-C., Mikou, M, Ferretti, C. (dir.) (2018), [Les dépenses de santé en 2017 – Résultats des Comptes de la santé – Édition 2018](#). Paris, France : DREES, coll. Panoramas de la DREES-santé.

Loiseau, R. (2020), [Aide au paiement d'une complémentaire santé : un niveau de couverture équivalent à celui des contrats du marché individuel en 2018](#). DREES, *Études et Résultats*, 1142.

Perronnin, M. (2019), [L'enquête Protection sociale complémentaire d'entreprise 2017](#). Irdes, *Les rapports de l'Irdes*, 572.

Annexe 1. Questionnaire couverture complémentaire de SRCV

Seules sont retranscrites ici les questions importantes pour la construction du modèle, et notamment pour la connaissance du statut de couverture dans SRCV. Des questions additionnelles portant sur l'ALD, les arrêts de travail, les raisons de la non-couverture ou encore le niveau de garanties sont posées dans l'enquête mais non retranscrites ici.

COUV **Actuellement, avez-vous [^PRENOM a-t-il(elle)] la CMU complémentaire ?³⁵³**
1. Oui
2. Non

➤ *Si non, NSP ou REFUS*

COMPSANT **Avez-vous [^PRENOM a-t-il(elle)] une complémentaire santé, appelée parfois aussi « mutuelle » qui complète les remboursements de la sécurité sociale ?³⁵⁴**
1. Oui
2. Non

353 La Couverture Maladie Universelle complémentaire appelée plus couramment CMU propose gratuitement une couverture en plus de la Sécurité Sociale pour les personnes dont les ressources sont faibles. Les bénéficiaires de la CMU complémentaire n'avancent pas d'argent quand ils vont chez le médecin.

354 Nous parlons ici de la complémentaire santé qui couvre le risque maladie en complément de l'assurance maladie, ce qui exclut les risques liés au décès, à l'incapacité, à l'invalidité et à la dépendance. Si l'enquêté a plusieurs contrats, les questions suivantes concerneront celui auquel les feuilles de soins de la personne sont le plus souvent transmises.

➤ *Si oui (COMPSANT=1)*

COUVENT **Avez-vous [^PRENOM a-t-il(elle)] obtenu cette complémentaire par l'entreprise dans laquelle vous[lui(elle)] ou un membre de votre[sa] famille travaille ou a travaillé ?³⁵⁵**
1. Oui
2. Non

RATTACH **Vous êtes [^PRENOM est] assuré(e) auprès de votre [cette] complémentaire santé ou mutuelle :**
1. En tant que souscripteur du contrat
2. En étant rattaché(e) à un contrat souscrit par un membre de votre [sa] famille

➤ *Si RATTACH=2, aller à la fin du bloc*

➤ *Si RATTACH=1 et COUVENT=2, NSP ou REFUS, aller à COMBPERS*

COMBPERS **Combien de personnes du ménage sont couvertes par cette complémentaire santé ?³⁵⁷**
Valeur numérique comprise entre 0 et 20

355 Les contrats souscrits auprès des organismes complémentaires sont, soit des contrats individuels, c'est-à-dire conclus directement par un individu auprès d'un organisme, soit des contrats collectifs, c'est-à-dire conclus par un employeur pour ses salariés. Rappel : si l'enquêté a plusieurs contrats, les questions concernent celui auquel les feuilles de soins de la personne sont le plus souvent transmises.

356 La complémentaire santé de base est la complémentaire la moins chère proposée par l'entreprise, offrant aussi le moins de garanties. Un niveau de garanties est « supérieur » dès lors que le salarié a souscrit une couverture plus élevée, en souscrivant une formule plus généreuse (aussi appelée option) ou un contrat offrant des remboursements additionnels (on parle alors de surcomplémentaire).

357 Inclure le souscripteur du contrat.

COMB **Combien déboursez-vous [^PRENOM débourse-t-il(elle)] pour votre [sa] complémentaire santé par mois ?³⁵⁸**

Valeur numérique comprise entre 0 et 5000

➤ *Si NSP ou REFUS à COMB*

COMBTR **Si vous ne pouvez pas ou ne souhaitez pas donner un montant précis du coût de votre [la] complémentaire santé [de ^PRENOM], à combien environ l'estimez-vous par mois ?**

1. À moins de 60 euros
2. De 60 à moins de 90 euros
3. De 90 à moins de 120 euros
4. De 120 euros à plus

³⁵⁸ Indiquez le coût total en euros pour l'ensemble des personnes rattachées à ce contrat de mutuelle, hors participation employeur dans le cas d'un contrat collectif. Tenir compte uniquement du coût de la complémentaire santé sans la prévoyance (ce qui exclut les risques liés au décès, à l'incapacité, à l'invalidité et à la dépendance). Lorsque le contrat est collectif, ce montant est en général prélevé directement par l'employeur et apparaît donc sur la fiche de paie du salarié, dans la colonne de montants retenus sur la rémunération du salarié. Il se peut que l'entreprise finance intégralement la couverture, auquel cas le montant à renseigner est de 0 euro.

Les dossiers de la DREES

N° 87 • décembre 2021

Le modèle de micro-simulation Ines-Omar 2017

Directeur de la publication

Fabrice LENGART

Responsable d'édition

Valérie BAUER-EUBRIET

ISSN

2495-120X

Ministère des Solidarités et de la Santé
Direction de la recherche, des études, de l'évaluation et des statistiques (DREES)

14 avenue Duquesne - 75 350 paris 07 SP
Retrouvez toutes nos publications sur drees.solidarites-sante.gouv.fr et nos données sur www.data.drees.sante.fr
