

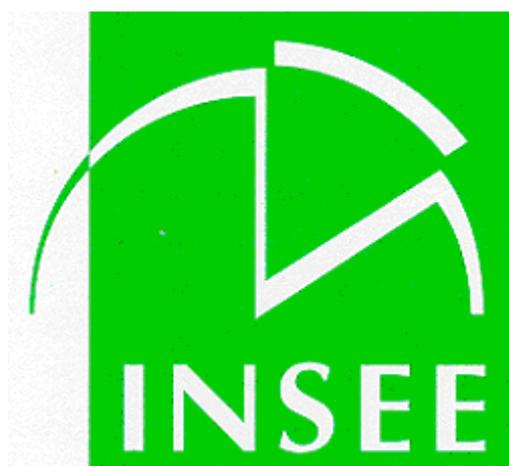
Direction des Statistiques Démographiques et Sociales

N° F1506

**Nowcasting du taux de pauvreté
par la micro-simulation**

Maëlle Fontaine, Juliette Fourcot

DOCUMENT DE TRAVAIL



Institut National de la Statistique et des Etudes Economiques

INSTITUT NATIONAL DE LA STATISTIQUE ET DES ÉTUDES ÉCONOMIQUES

Série des documents de travail
de la
DIRECTION DES STATISTIQUES DÉMOGRAPHIQUES ET SOCIALES

N° F1506

Nowcasting du taux de pauvreté par la micro-simulation

Auteurs : Maëlle FONTAINE et Juliette FOURCOT
(Division des Études sociales)

Document de travail

décembre 2015

*Ces documents de travail ne reflètent pas la position l'Insee et n'engagent que leurs auteurs.
Working papers do not reflect the position of Insee but only their author's views.*

Nowcasting du taux de pauvreté par la micro-simulation

Résumé : Chaque année, l'Insee publie en septembre **N+2** les principaux indicateurs de pauvreté et de niveau de vie relatifs à l'année **N**. Ce délai n'est pas satisfaisant pour répondre à la demande sociale des utilisateurs de ces indicateurs (médias, décideurs publics). Ces 21 mois entre la fin de l'année considérée et la publication du taux de pauvreté sont pour les trois quarts environ dus au temps de recueil des données fiscales et sociales, et pour un quart environ liés aux appariements statistiques qui permettent de produire l'enquête Revenus Fiscaux et Sociaux (ERFS). L'exercice de nowcasting consiste à produire plus précocement (à l'automne **N+1**) un indicateur du taux de pauvreté de l'année **N**. La méthode envisagée pour cet exercice est celle de la micro-simulation, qui reconstitue le niveau de vie des individus en imputant sur barèmes les prestations et les prélèvements, ce qui permet de tenir compte de modifications légales de ces dispositifs. L'exercice est basé sur le modèle Ines (Insee-Drees), qui permet de simuler la majorité des législations socio-fiscales françaises à partir de 1990, sur la base de n'importe quel millésime de l'ERFS. Si la micro-simulation rend possible cet objectif, il reste à mesurer la qualité des indicateurs précoces ainsi produits. L'enjeu de ce document de travail est donc de comparer, sur des périodes passées, les résultats qui auraient été délivrés par la micro-simulation à ceux qui ont réellement été diffusés à partir de l'ERFS, avant de se prononcer sur la pertinence du nowcasting par micro-simulation. Pour ce travail, des informations fiscales qui n'auraient pas été disponibles en conditions réelles ont été utilisées.

Les premiers résultats de l'exercice de *nowcasting* réalisé sur trois points (niveaux de vie de 2009 à 2012) sont satisfaisants, puisque le sens de variation de l'indicateur principal, le taux de pauvreté relatif au seuil de 60 % du niveau de vie médian est correctement « nowcasté » par Ines. En revanche, l'évolution du taux de pauvreté à 50 % n'aurait pas toujours été estimée dans le bon sens selon les années. Néanmoins, compte tenu de la difficulté à produire une telle projection et la forte sensibilité d'un tel indicateur aux méthodes de simulation des transferts sociaux, il est recommandé de considérer notre estimateur nowcasté avec grande précaution.

Mots-clé : micro-simulation, *estimation avancée*, pauvreté, inégalités

Codes JEL : C81, I32, D31, H31

Nowcasting of poverty rate by microsimulation

Abstract : Every year in September **N+2**, INSEE publishes the main poverty and standard of living indicators for year **N**. This delay is unsatisfactory for meeting the social requirements of those who use these indicators (media, public decision-makers). Of the 21 months between the end of the year under consideration and the publication of the poverty rate, about three-quarters of this time is taken up collecting tax and social data, and about one quarter with statistically matching from which the Tax and Social Incomes Survey (ERFS) is produced. The practice of nowcasting consists of producing an earlier indicator of the poverty rate for year **N** (in autumn **N+1**). The method to be used for this exercise is microsimulation, which creates individuals' standard of living by imputing benefits and contributions on scales, and thus it is possible to take account of any legal changes made to these measures. The exercise is based on the INES (INSEE-DREES) model, which simulates the majority of French social security and tax legislation from 1990, based on any year of the ERFS. Although in technical terms, microsimulation can meet this objective, we must now measure the quality of the early indicators thus produced. The purpose of this working document is therefore to look at periods in the past, and compare the results that would have been produced by microsimulation with those that were in fact disseminated from the ERFS, before reaching a decision on the relevance of nowcasting by microsimulation. In this document, some fiscal information that would not have been available in real conditions, has been used.

The first results of the nowcasting exercise on standards of living from 2010 to 2012 are satisfactory, since the direction of variation in the main indicator, the poverty rate in relation to 60% of the median standard of living, is correctly “nowcasted” by INES. Nevertheless, given the difficulty in producing projections and high sensibility of such an indicator to the microsimulation method affecting social transfers, our nowcasted estimator must be considered with great caution.

Key words: micro-simulation, nowcasting, timeliness, poverty, inequalities

Classification JEL : C81, I32, D31, H31

Table des matières

1. Principe du nowcasting.....	7
1.1. Enjeux.....	7
1.2. Méthodes.....	8
1.2.1. Méthode par régression de quantiles non conditionnels.....	8
1.2.2. Méthode par micro-simulation.....	9
2. Précédents exercices de nowcasting par micro-simulation.....	12
2.1. Travaux divers menés avec le modèle Ines (2004, 2008, 2009, 2012).....	12
2.2. Travaux de nowcasting avec Euromod en 2013.....	12
3. Méthode.....	15
3.1. Rappels sur le fonctionnement habituel d'Ines.....	15
3.2. Détourner Ines pour faire du nowcasting.....	17
3.3. Définition des estimateurs.....	19
3.4. Différences conceptuelles de niveau de vie entre la microsimulation et l'appariement.....	21
4. Résultats.....	23
4.1. Comparaison entre les résultats de l'ERFS et l'évaluation contemporaine.....	23
4.1.1. Estimation du niveau de vie et de ses composantes.....	23
4.1.2. Indicateurs de pauvreté et d'inégalités.....	25
4.2. Comparaison entre les résultats de l'ERFS et ceux de l'évaluation décalée.....	26
4.2.1. Estimation du niveau de vie et de ses composantes.....	27
4.2.2. Estimation des indicateurs de pauvreté et d'inégalités.....	28
4.3. Nowcasting avec mise en conditions réelles.....	28
5. Conclusion et enjeux.....	31
6. Bibliographie.....	32
Annexes.....	33

1. Principe du nowcasting

1.1. Enjeux

En septembre **N+2** de chaque année, l'Insee publie les premiers résultats du nouveau millésime de l'Enquête Revenus Sociaux et Fiscaux (ERFS) sur les niveaux de vie en **N** [HOU14]. À cette occasion, est rendu public le taux de pauvreté de l'année **N**, indicateur particulièrement attendu du grand public et des décideurs et largement repris dans les médias. C'est donc 21 mois après la fin de la période considérée que les premiers résultats sur les niveaux de vie sont diffusés. Ce délai, qui peut paraître long par rapport aux attentes des utilisateurs, est essentiellement dû aux spécificités du système d'information. En effet, l'ERFS¹ est le résultat de l'appariement statistique entre l'Enquête Emploi, les fichiers fiscaux de la Direction Générale des Finances Publiques (déclaration de revenu et taxe d'habitation) et les fichiers administratifs de la Caisse Nationale des Allocations Familiales (Cnaf), de la Caisse Centrale de la Mutualité Sociale Agricole (CC-MSA) et de la Caisse Nationale d'Assurance Vieillesse (Cnav). Le calendrier de production de l'ERFS est le suivant : les informations fiscales et les fichiers sociaux² sont respectivement transmis 12 mois et 15 mois après la fin de la période considérée (décembre **N+1** et avril **N+2**) ; une fois ces données reçues, l'Insee procède à un appariement statistique qui aboutit à la fin du mois de juin **N+2**. Les deux mois suivants sont consacrés aux traitements statistiques pour produire les indicateurs clés de pauvreté et d'inégalités.

L'appariement statistique avec les déclarations fiscales et les fichiers sociaux apparaît dès lors comme une contrainte forte sur les délais de production. Une enquête directe auprès des ménages sur leurs revenus pourrait en principe donner des résultats plus précoces, mais le calendrier exact mériterait néanmoins d'être analysé de près³. Toutefois, le questionnaire doit être suffisamment détaillé sur les revenus pour obtenir un niveau de vie satisfaisant, ce qui allonge rapidement les temps d'interrogation. Un tel dispositif a existé jusqu'au début des années 2000, mais il a été abandonné au profit d'une méthode d'appariement statistique avec des données administratives, produisant des résultats de meilleure qualité⁴ et moins coûteuse qu'une enquête, mais exigeant en revanche des délais plus longs.

La demande à raccourcir les délais de production de cet indicateur phare est forte, en France mais aussi de la part d'Eurostat qui souhaiterait disposer d'indicateurs plus précocement de manière à orienter la politique européenne de lutte contre la pauvreté. Cette demande est très bien illustrée par le fait que l'un des cinq objectifs de la stratégie Europe 2020 pour une croissance intelligente, durable et inclusive, adoptée par le Conseil Européen en juin 2010, est précisément de réduire d'un sixième l'indicateur européen de pauvreté entre 2007 et 2017. Le gouvernement français a décliné à l'échelle nationale cet objectif, lorsque le Président de la République a annoncé le 17 octobre 2007 le projet de faire baisser d'un tiers de la pauvreté en 5 ans. À défaut de pouvoir réduire considérablement le délai de production des niveaux de vie, pour tenter malgré tout de répondre à cette demande d'indicateurs précoces pour le suivi des politiques de lutte contre la pauvreté, deux méthodes ont été explorées à la Direction des

1 L'ERFS existe sous cette forme depuis 2006. Auparavant, la base de référence pour le calcul des niveaux de vie était réalisée sans appariement avec les fichiers sociaux (Enquête Revenus Fiscaux), et encore avant, uniquement à partir des déclarations fiscales.

2 Les fichiers de la CC-MSA sont reçus plus tôt que les fichiers CNAV et CNAF, en octobre **N+1**.

3 Pour un même volume d'enquête (et donc de précision) que l'ERFS actuelle, soit 60 000 ménages, une période de collecte de trois à quatre mois serait nécessaire, soit au mieux une fin de collecte fin avril de l'année **N+1**. L'aval informatique prend ensuite environ un mois, soit des données disponibles pour les équipes statistiques fin mai de l'année **N+1**. Resteraient ensuite quatre mois pour faire mieux que ce qui est proposé avec la microsimulation, compte tenu que l'on sait que les données de revenus collectées en face à face sont moins bien renseignées, que les montants sont sujets à des valeurs manquantes et des phénomènes d'arrondis qu'il est nécessaire de redresser par imputation stochastique. Enfin, le coût terrain d'une telle opération se chiffrerait à plusieurs millions d'euros.

4 L'utilisation des déclarations de revenus pour le calcul des niveaux de vie présente toutefois d'autres difficultés, comme l'imputation de revenus défiscalisés (livret A, livret de développement durable, PEL...), par nature absents des déclarations.

Statistiques Démographiques et Sociales pour produire des indicateurs d'évolution. Cet exercice porte le nom de *nowcasting* en anglais, par analogie du terme *forecasting*, puisqu'il s'agit de projeter sur une période déjà révolue. La division Études Sociales a développé une méthode de *nowcasting* par microsimulation, qui est présentée dans cette note, tandis que la division Revenus et Patrimoine des ménages a exploré les possibilités ouvertes par des méthodes de régression de quantiles non conditionnels, dont le principe et les résultats seront brièvement rappelés. Ces indicateurs simulés (soit par micro-simulation, soit économétriquement) dépendent fortement de la capacité à bien appréhender des évolutions de distribution qui ne sont pas disponibles au moment où l'on fait le *nowcasting*, et risquent d'être remis en cause par la suite par les résultats définitifs issus de l'ERFS. Le risque d'erreur constaté *a posteriori* sur les prévisions est d'autant plus grand en période de retournement conjoncturel, situation dans laquelle l'erreur est justement la plus préjudiciable. Des améliorations pourraient être apportées à la technique de microsimulation présentée ici, en jouant notamment sur les pondérations pour tenir compte de la déformation de la population des chômeurs en termes de catégorie sociale⁵.

1.2. Méthodes

1.2.1. Méthode par régression de quantiles non conditionnels

Les résultats sont issus de travaux effectués par deux stagiaires⁶ accueillis en 2012 et 2013 à la division Revenus et Patrimoine des Ménages. Les méthodes sont appliquées à l'enquête Statistiques sur les Ressources et les Conditions de Vie (SRCV), composée de 12 000 ménages, en panel rotatif sur 9 ans. Depuis SRCV 2008 (sur les revenus 2007), les revenus sont issus d'un appariement avec les sources administratives, qui suit une méthode proche de celle mise en œuvre pour l'ERFS. Toutefois, le questionnaire comporte une question très simple sur le « revenu mensuel total du mois précédant la collecte », pour l'ensemble du ménage.

Le principe des méthodes de prédiction à partir de modèles de régression consiste à estimer sur les données passées une relation économétrique entre le revenu « administratif » (ou une transformée Y_{admin} de ce revenu) et un ensemble de variables explicatives (X), incluant le revenu « collecté » ($Y_{collecte}$) et les revenus « administratifs » antérieurs pour les ménages réinterrogés. Le modèle le plus simple auquel on peut penser est un modèle de régression linéaire :

$$Y_{admin} = a * Y_{collecte} + b * X + u$$

Compte tenu de notre objectif de prédire un taux de pauvreté, un autre modèle assez simple consiste à modéliser le fait d'être pauvre ou non, selon un modèle logistique (Logit ou Probit dichotomique en fonction de la loi des résidus choisis) sur la variable latente z^* .

$$P(pauvre_{admin}) = P(z^* < 0)$$

$$\text{avec } z^* = a * Y_{collecte} + b * X + \varepsilon$$

Toutefois, si ce deuxième modèle est bien adapté à la prédiction d'une variable binaire, il ne permet pas de traiter facilement la prédiction d'autres indicateurs d'inégalités ou de pauvreté. C'est ici que les méthodes de régression de quantiles non conditionnels prennent leur sens car elles permettent de modéliser n'importe quel résumé de la distribution et donc d'appliquer une

5 La méthode actuelle garantit par calage un nombre total de chômeurs et un nombre total d'individus par catégorie sociale correspondant à la réalité contemporaine, mais ces deux dimensions ne sont pas croisées. Les effets de composition éventuels au sein de la population des chômeurs et leurs conséquences sur les différentes composantes du niveau de vie ne sont donc pas pris en compte (si entre deux dates, le nombre de chômeurs est inchangé, mais que la population des chômeurs se compose désormais d'une majorité de cadres qui ont perdu leur emploi, alors au niveau macro, les revenus salariaux baissent et les revenus du chômage, calculés sur la base des salaires antérieurs, augmentent).

6 Lukas Kuld (2012), Pierre Pora (2013).

méthode identique pour la prédiction de n'importe quel indicateur usuel d'inégalités (taux de pauvreté, rapport interdécile, rapport des masses 100-S80/S20, indice de Gini...). Le principe est de régresser une transformée de la variable d'intérêt, la *Recentered Influence Function* (RIF), sur les variables explicatives :

$$RIF_{\text{indicateur}}(Y_{\text{admin}}) = a * RIF_{\text{indicateur}}(Y_{\text{collecte}}) + b * X + u$$

La RIF est spécifique à chaque indicateur que l'on souhaite prédire. La somme des valeurs individuelles prédites de la RIF donne ensuite la prédiction souhaitée de l'indicateur.

Les travaux de Pora (2013) retiennent cinq modèles différents :

- deux modèles binaires du fait d'être pauvre ou non (Logit dichotomique et Probit dichotomique) ;
- deux modèles linéaires généralisés pour la RIF du taux de pauvreté ;
- un modèle de type logit multinomial pour la RIF du taux de pauvreté, car dans le cas d'un taux de pauvreté relatif, cette fonction prend trois valeurs non monotones (une pour les individus sous le seuil de pauvreté, une pour les individus entre le seuil et la médiane des niveaux de vie, une dernière pour les individus au-dessus de la médiane).

Modèle	Taux 2009	Taux 2010	Variation 2009-2010	Sens
EU-SILC	12,9	13,3	0,4	+
Modèle linéaire généralisé	12,8	14,5	1,7	+
Modèle linéaire généralisé (réduit)	13,1	14,7	1,6	+
Logit dichotomique	12,6	13,8	1,2	+
Probit dichotomique	12,6	13,7	1,1	+
Logit multinomial	11,2	12,8	1,6	+

Tableau 1 : Taux de pauvreté à 60 % issus des différents modèles envisagés (résultats issus du rapport de stage de Pierre Pora).

L'intérêt de l'application de la méthode à l'enquête SRCV était de profiter, pour l'étape de prédiction, de l'information contemporaine sur l'évolution de la distribution du revenu mensuel collecté en face à face. Toutefois, les données collectées sont affectées de nombreux arrondis et valeurs manquantes qui pénalisent vraisemblablement la qualité de l'ajustement. Aucun modèle n'est réellement satisfaisant et le modèle linéaire généralisé (régression quantile) est plutôt moins bon en termes de prédiction qu'un simple probit dichotomique (tableau 1). Au final, tous ces modèles économétriques surestiment l'évolution du taux de pauvreté d'au moins 0,7 point (la prévision allant de +1,1 à +1,7 point au lieu de +0,4 point).

Parmi les limites des travaux actuels, on peut retenir d'une part l'insuffisante utilisation de l'information longitudinale antérieure pour les ménages réinterrogés, et d'autre part la complexité à prédire la RIF du taux de pauvreté⁷.

1.2.2. Méthode par micro-simulation

La micro-simulation du système social et fiscal consiste à calculer sur barèmes des prestations et prélèvements en utilisant des données représentatives de la population d'intérêt. Ainsi, partant des informations disponibles sur les différents revenus, il est possible de calculer les droits ouverts aux prestations familiales, les minima sociaux, les allocations logement d'une part, et d'autre part l'impôt sur le revenu, les prélèvements sociaux, les cotisations et contributions

7 Un modèle linéaire ne paraît pas vraiment adapté compte-tenu du caractère multinomial non monotone de la fonction, mais l'utilisation d'un logit multinomial en prédiction et pas seulement en estimation pose également problème.

sociales⁸. En décomptant ces éléments du revenu brut, on peut reconstruire le revenu disponible du ménage et en déduire son niveau de vie⁹.

Plusieurs modèles de micro-simulation sont actuellement utilisés pour calculer des niveaux de vie. À l'exception d'OpenFisca, ils sont conçus pour être utilisés sur une seule source statistique. Trois modèles français s'appuient sur l'ERFS : Ines (Insee-Drees), Myriade (Cnaf) et Saphir (Direction du Trésor). Taxlpp (Institut des Politiques Publiques) s'appuie sur un échantillon simulé à partir de sources diverses¹⁰ et OpenFisca (France Stratégie) sur des cas-types à défaut de pouvoir accéder à une source représentative de revenus. Le modèle Euromod de l'Union Européenne, développé à l'Institute for Social and Economic Research (Iser) s'appuie, lui, sur la version européenne de SRCV (EU-SILC). À notre connaissance, seuls Ines et Euromod ont été utilisés pour produire un indicateur précoce du taux de pauvreté en France.

À partir d'un millésime d'enquête donné, ERFS ou SRCV, tous ces modèles se fondent sur des projections à des horizons divers, pour répondre à différents besoins, le plus souvent dresser l'état des lieux d'un ou plusieurs systèmes (comparaisons internationales, analyses d'ensemble du système redistributif) ou établir des analyses en variante (études *ex-post* ou *ex-ante* d'impact de nouvelles mesures fiscales ou sociales, comparaison de différentes années de législation...). Ces modèles sont donc tous potentiellement capables de produire des taux de pauvreté précoces par rapport à ceux de l'ERFS et de SRCV. Dans le cas des modèles Ines et Myriade, les projections sont faites sur 2 ans. Ils s'appuient ainsi par exemple sur l'ERFS 2011 pour simuler des niveaux de vie 2013. Comme les rebasages de ces modèles (adaptation du modèle à l'ERFS suivante) sont faits au cours du printemps **N+3** (printemps 2014 sur l'ERFS 2011), les projections concernent la dernière année révolue (**N+2**) (voir illustration 1). Le modèle Saphir se fonde quant à lui à proprement parler sur la projection d'une période future, avec pour objectif de réaliser des chiffrages pour les projets de loi de finance. Le modèle Euromod est généralement utilisé en variante, à l'exception des travaux récents de *nowcasting*.

Dans cette étape de projection pour micro-simuler les niveaux de vie, la principale difficulté technique consiste à obtenir une structure socio-démographique et des revenus qui ne seront pas simulés sur barème plausible (l'actualisation des prestations et prélèvements ne constitue pas une difficulté majeure à ce stade puisqu'il « suffit » de mettre à jour les règles de calcul et les barèmes selon la législation en vigueur l'année étudiée). On distingue deux grandes méthodes pour projeter l'échantillon représentatif :

- La première et la plus naturelle est de simuler des trajectoires individuelles. Tous les individus vieillissent d'une (ou plusieurs) année(s), certains meurent, d'autres ont des enfants, trouvent un emploi ou le perdent, etc. Aussi simple que cette méthode puisse paraître, il s'agit d'utiliser des probabilités de transitions entre les différents états, ce qui peut devenir rapidement complexe si l'on souhaite refléter la diversité des situations (mortalité par âge, transitions emploi-chômage en fonction de l'âge, du secteur et de la catégorie socioprofessionnelle, estimations du salaire pour les personnes trouvant un emploi). C'est toutefois la méthode adoptée par Euromod dans ses exercices de *nowcasting*, sans naissance ni décès, mais uniquement avec des destructions et créations d'emplois au niveau individuel. On appelle cette méthode de projection le vieillissement dynamique.

8 Les allocations chômage et pensions de retraite, bien qu'il s'agisse de transferts sociaux calculés sur barème, sont généralement exclues du champ des transferts lors d'analyses redistributives en coupe. Deux raisons à cela : d'une part, elles sont fortement associées à une logique assurantielle dont la redistributivité s'analyse mieux en cycle de vie, et d'autre part, leur calcul requiert une connaissance fine de la trajectoire professionnelle sur une durée supérieure à un an ce qui n'est généralement pas présent dans les données habituellement utilisées.

9 Notons que lorsque l'on s'appuie sur des données fiscales, on s'intéresse à des revenus déclarés, nets de cotisations. Dans Ines en revanche, les cotisations sont simulées à partir des revenus déclarés puis réduites du revenu brut ainsi calculé, afin de pouvoir mesurer l'impact des cotisations en termes de redistribution.

10 Ces sources sont détaillées dans le document méthodologique de Taxlpp [BOZ12].

- La deuxième méthode, adoptée par l'ensemble des modèles français, consiste à modifier les pondérations pour refléter une structure légèrement déformée, plus récente. Il s'agit d'une utilisation originale des algorithmes de calage sur marges, habituellement utilisés pour redresser les pondérations d'une enquête. L'avantage de cette méthode est qu'elle est simple à mettre en œuvre. En revanche, elle n'est adaptée que pour de faibles modifications de structure, et ne conviendrait probablement pas lorsque des modifications économiques d'ampleur ont lieu : impact fort d'une crise économique comme en Grèce, introduction de dispositifs modifiant considérablement la structure de l'emploi (par exemple introduction d'un salaire minimum), fortes migrations de population. Comme les individus ne vieillissent pas individuellement, mais qu'on fait plutôt l'hypothèse d'une déformation globale de la population sur quelques marges, on parle de vieillissement statique.

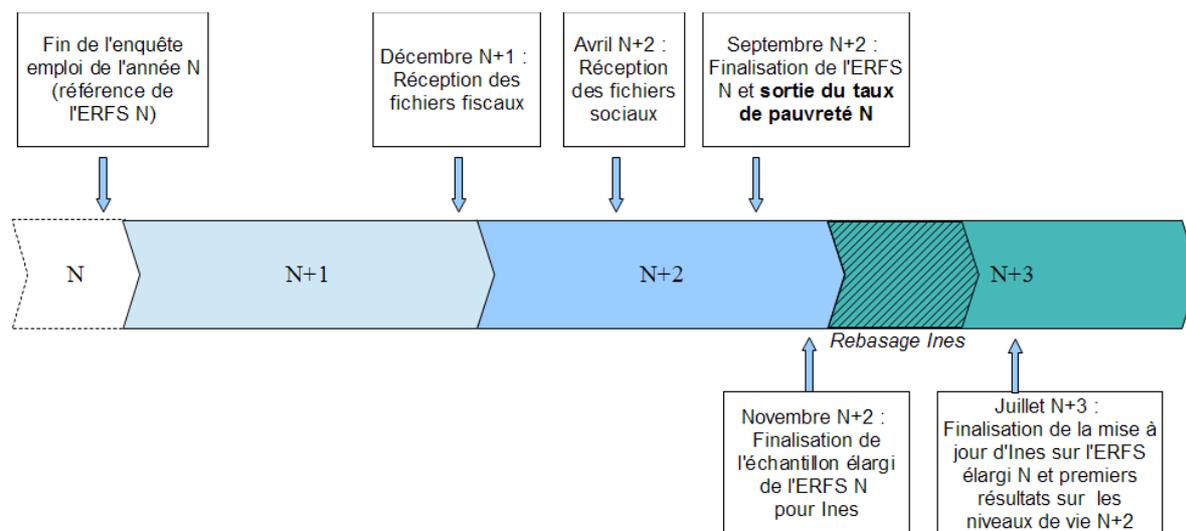


Illustration 1: Déroulement du processus de production de l'ERFS (en haut) et d'Ines (en bas).

L'une et l'autre de ces deux méthodes projetant les niveaux de vie des ménages sur plusieurs années rendent possible de calculer des indicateurs distributionnels plus précoces que ceux produits habituellement, comme le taux de pauvreté, mais aussi le rapport inter-déciles, le coefficient de Gini, etc. Certaines différences méthodologiques entre les sources (ERFS ou SRCV) et les modèles de micro-simulation induisent structurellement des écarts de niveaux entre les indicateurs tirés des sources et ceux produits par les modèles. Cependant, sous l'hypothèse que ces écarts de niveaux sont constants d'une année sur l'autre, il est possible de comparer les évolutions des indicateurs tirés des sources et produits par les modèles. C'est pourquoi dans la partie 4 (résultats), nous raisonnerons en évolution pour neutraliser les écarts de méthodologie. Par ailleurs, il est évident que les indicateurs statistiques fournis sont de moins en moins précis et fiables à mesure que l'on s'approche de la période contemporaine puisque les informations pour vieillir les données sont de plus en plus frustes.

2. Précédents exercices de nowcasting par micro-simulation

2.1. Travaux divers menés avec le modèle Ines (2004, 2008, 2009, 2012)

Par le passé, le modèle Ines a déjà été mobilisé à différentes reprises pour des exercices de *nowcasting* ou de *forecasting* du taux de pauvreté, dont les méthodes et les principaux résultats peuvent être rappelés utilement ici (tableau 2). Dans l'ensemble, ces exercices se sont révélés *a posteriori* non concluants. Il faut noter cependant qu'ils ont été menés avec des objectifs divers (réponse à la demande parlementaire ou ministérielle, bilan technique du modèle Ines...), mais toujours différents de celui qui est le nôtre ici (production d'un indicateur avancé du taux de pauvreté). Ces exercices sont présentés en détail en annexes 1 à 4.

	Drees 2004	Drees 2008	Drees 2009	Insee 2012
<i>Horizon de prévision</i>	2002	2006-2011	2007-2009	2011
<i>Données d'origine</i>	ERF 1999	ERFS 2004	ERFS 2006	ERFS 2009
<i>Méthode de vieillissement</i>	dérive sur deux ans + inflation sur un an	dérive sur deux ans + scénarios macroéconomiques	dérive sur deux ans	dérive sur deux ans
<i>Données de comparaison</i>	ERF 2001 ou ERF 2002	ERFS 2006 à 2011	ERFS 2007 à 2009	ERFS 2011
<i>Mise en commun du champ</i>	aucune	aucune	aucune	- travail sur l'ERFS noyau - retrait des transferts en nature du calcul du revenu disponible

Tableau 2: Résumé des précédents exercices de nowcasting du taux de pauvreté.

L'analyse conduite ici se distinguera de tous ces travaux antérieurs par la mise en œuvre d'une « dérive inverse » plus fine et par un effort plus précis pour faire converger la méthodologie d'Ines vers celle de l'ERFS. Une fois ces différences réduites, nous nous donnerons comme critère de comparer si les évolutions du taux de pauvreté *nowcastées* par Ines sont cohérentes avec celles produites par l'ERFS.

2.2. Travaux de nowcasting avec Euromod en 2013

Euromod est le modèle de micro-simulation de l'ensemble des pays de l'Union Européenne. Géré par Eurostat, il s'appuie sur l'échantillon représentatif du dispositif *European Union – Statistics on Income and Living Conditions* (EU-SILC). En 2012 puis en 2013, Euromod a mené deux exercices de *nowcasting* et de *forecasting* des taux de pauvreté annuels, d'abord sur la période 2008-2012 à partir d'EU-SILC 2008 (revenus 2007), puis sur la période 2010-2013 à partir d'EU-SILC 2010 (revenus 2009) [NAV13][LEV13]. La première étude présentait les résultats pour 9

pays alors que la deuxième traite de 13 États de l'Union Européenne (France, Allemagne, Autriche, Finlande, Italie, Espagne, Portugal, Pologne, Grèce, Estonie, Lettonie, Lituanie et Roumanie), dont certains connaissent des conditions économiques très volatiles pendant la crise. Les résultats récapitulés ci-après sont ceux obtenus à partir d'EU-SILC 2010 pour la période 2010-2013.

À la différence du modèle Ines, Euromod est un modèle dynamique (cf. 1.2.2). Autrement dit, il met à jour la structure du marché du travail via des simulations de transitions individuelles aléatoires entre emploi, chômage et chômage de longue durée, qui dépendent de trois paramètres (sexe, âge, niveau de diplôme). Le modèle de transitions est ainsi estimé sur 18 strates, à partir d'EU-SILC, et les niveaux d'emploi et de chômage sont calés sur les résultats des *Labor Force Survey* (LFS), enquêtes également harmonisées au niveau européen.

Les revenus qui ne sont pas par la suite calculés sur barèmes sont ensuite mis à jour. Pour les années passées, ce sont des données conjoncturelles issues d'autres enquêtes ou des données administratives qui sont utilisées. Pour l'année en cours, ce sont des prévisions macroéconomiques officielles ou, à défaut, des estimations avec les données trimestrielles ou des indicateurs tels que l'inflation, qui sont utilisés. De plus, les transferts sont simulés selon la législation en vigueur au 30 juin de chaque année. Enfin, pour les individus qui changent de position sur le marché du travail, ceux qui trouvent un emploi obtiennent le salaire moyen de leur strate et ceux qui deviennent chômeurs se voient calculer leurs droits selon des règles spécifiques à chaque pays.

Pour compléter la description de la méthodologie de cette étude, il faut ajouter qu'un calage est réalisé au niveau de chaque ménage. Un coefficient additif est en effet appliqué, de manière à ce que le niveau de vie simulé du ménage en 2009 corresponde à celui calculé par EU-SILC 2010 (revenus 2009, variable HX090). Par construction, les taux de pauvreté coïncident donc pour 2009. Seule la France fait exception, avec un écart de 0,2 point en 2009, car la version de EU-SILC utilisée dans le modèle (SILC 2010-1) diffère de celle utilisée comme référence de comparaison (SILC 2010-5). La cale est ensuite reportée au même niveau pour les années ultérieures.

Euromod présente l'avantage de la comparabilité des différents États analysés, permise par l'utilisation des enquêtes SILC et LFS, toutes deux harmonisées au niveau européen. Toutefois, dans le détail des dispositifs, le modèle peut s'avérer très imprécis, notamment pour la France où les masses de RSA versées sont trois fois plus élevées dans Euromod que dans des données-externes. De même, toujours pour la France, l'impôt sur le revenu est plus élevé d'environ 13 %, et les allocations chômage sont environ 2 fois plus élevées [AVR13].

	Allemagne	Autriche	Espagne	Estonie	Finlande	France	Grèce	Italie	Lettonie	Lituanie	Pologne	Portugal	Roumanie
2009	0,0	0,0	0,0	0,0	0,0	0,2	0,0	0,0	0,0	0,0	0,0	0,0	0,0
2010	0,0	-0,2	-1,2	-1,3	-0,7	0,6	-0,5	-1,3	0,8	2,2	-0,2	-0,6	-1,5
2011	-0,3	-2,1	-1,1	-0,5	-0,7	0,6	-1,3	-1,0	1,1	2,9	0,5	0,2	-2,5

Tableau 3 : Différence entre le taux de pauvreté à 60 % moyen estimé par EUROMOD à partir de EU-SILC 2010-1 et de micro-simulation dynamique, et les taux de pauvreté obtenus avec EU-SILC pour les années 2009 à 2011.

Note : pour la France, le faible écart sur le niveau de vie 2009 s'explique par le fait qu'Euromod s'appuie sur la première version de SILC 2010 alors que le point de référence est donné par la cinquième version du même millésime. [LEV13]

La comparaison des taux de pauvreté à 60 % d'Euromod avec ceux de EU-SILC conduisent aux conclusions suivantes :

- pour l'année 2010, c'est-à-dire en projection sur un an, l'écart est très faible (inférieur à 0,2 point) pour l'Allemagne, l'Autriche et la Pologne, modéré (compris entre 0,5 et 1 point) pour la Finlande, la France, la Grèce, la Lettonie et le Portugal, dont certains ont expérimenté une crise économique d'ampleur. L'écart est important (supérieur à 1 point) pour les 4 pays restants des 13 étudiés. L'écart absolu augmente ensuite à mesure que

l'on se rapproche de la période contemporaine. Le sens d'évolution du taux de pauvreté n'est pas toujours bien prédit ;

- l'utilisation du coefficient additif (la cale) conduit *de facto* à rapprocher les niveaux des taux de pauvreté projetés et observés (la réduction des écarts pouvant atteindre 3 à 4 points de pourcentage pour certains pays, cf. illustration 2). Même si la méthode n'assure pas de conserver exactement les mêmes évolutions d'une année à l'autre avant et après calage, les évolutions restent peu différentes. Il apparaît alors aussi légitime que transparent d'utiliser le modèle en évolution.

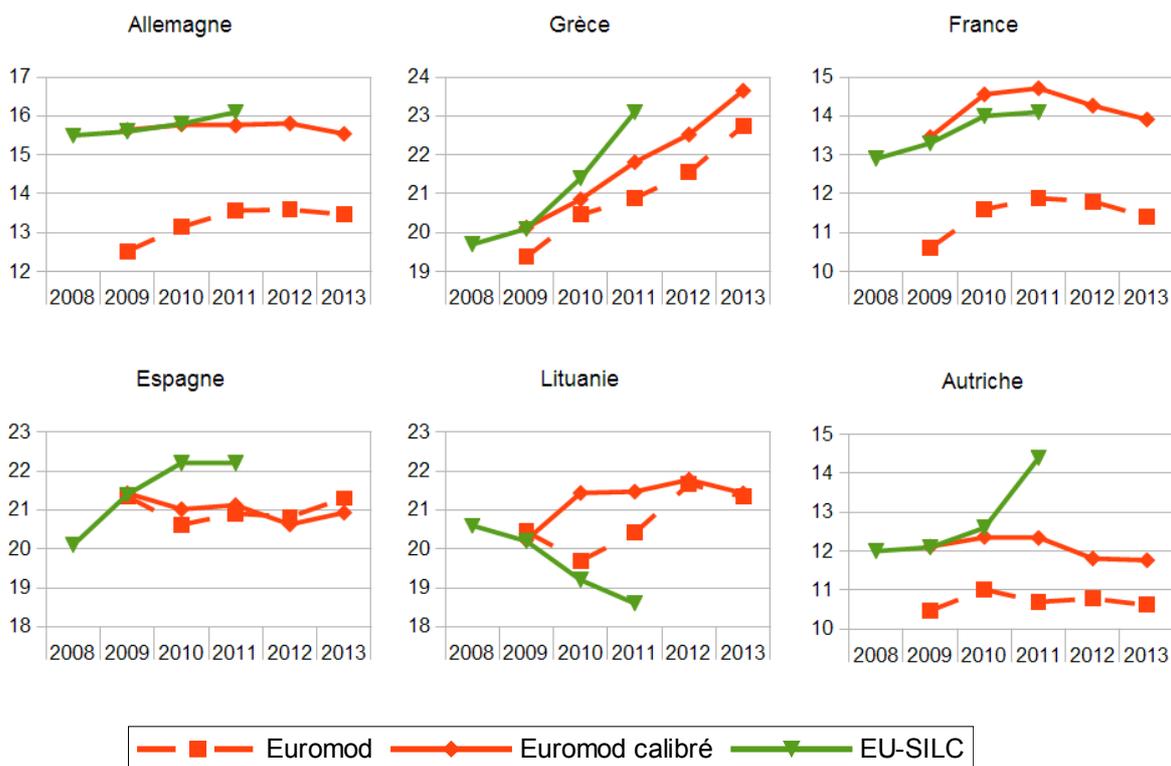


Illustration 2: Taux de pauvreté à 60 % estimés par Euromod avec et sans calibrage et par EU-SILC (2008 à 2011) [LEV13].

En conclusion de tous ces travaux, le *nowcasting* apparaît comme un exercice périlleux qui ne saurait se résumer en la simple production d'indicateurs à comparer avec des cibles issues d'autres sources. Aussi la maîtrise des facteurs potentiels de différences est cruciale pour une analyse judicieuse des résultats. En particulier, plusieurs leçons peuvent être tirées. Premièrement, les différents dispositifs doivent être simulés suffisamment finement pour s'assurer qu'une convergence du taux de pauvreté n'est pas simplement le fruit de compensations entre des transferts sous-estimés et d'autres sur-estimés (cas d'Euromod pour la France). Deuxièmement, il est important de travailler en amont au rapprochement des concepts de niveau de vie entre les données et le modèle, afin de ne pas imputer l'écart observé à des différences de champ facilement corrigeables (partie 3.3). Enfin, un soin tout particulier doit être placé sur le vieillissement des données utilisées, dont la méthodologie reste le point-clé de tout exercice de prévision (partie 3.2).

3. Méthode

3.1. Rappels sur le fonctionnement habituel d'Ines

Le modèle de micro-simulation Ines permet de simuler sur barème la grande majorité des prélèvements et prestations sociales intervenant dans le calcul du niveau de vie. Jusqu'en 2015, il s'appuie sur la version élargie de l'ERFS (par opposition à l'ERFS dite « noyau »). Dans son utilisation habituelle, le modèle sert à calculer au niveau individuel le revenu disponible en **N+2** (année révolue la plus récente au moment de la mise à jour) en s'appuyant sur le millésime **N** de l'ERFS. La dernière mise à jour du modèle a été achevée à la fin de l'été 2014 et les premiers résultats sur les niveaux de vie en 2013 ont été produits à partir de l'ERFS 2011 (voir figure 1).

La microsimulation effectuée dans Ines consiste, dans un premier temps, à vieillir la base de données de manière statique (modification de la structure socio-démographique et évolution de revenus qui ne dépendent pas des barèmes) et, dans un second temps, à simuler sur barème les prestations et prélèvements entrant en jeu dans le niveau de vie du ménage (sans utiliser les informations issues de l'appariement avec les fichiers sociaux et fiscaux). Dans la suite de cette partie, seule la première de ces deux étapes est détaillée, la seconde ne soulevant pas de question méthodologique importante.

Une spécificité de la législation française, prise en compte dans le modèle de micro-simulation Ines, est de faire intervenir la situation et en particulier les revenus des années antérieures dans le calcul de certaines prestations sociales¹¹ et prélèvements. L'impôt sur le revenu payé une année donnée (**N**) se calcule à partir des revenus et de la situation de l'année précédente (**N-1**). Les allocations logement sont calculées sur la base des revenus de l'antépénultième année (**N-2**). Comme le vieillissement statique génère *in fine* trois années consécutives de revenus pour chaque individu, cette particularité de la législation française est facilement prise en compte dans le modèle.

Dans son utilisation habituelle, les revenus de l'antépénultième année sont les revenus de l'ERFS. Ceux de l'année courante et de l'année antérieure sont obtenus par vieillissement statique (cf. partie 1.2.2), que l'on appelle aussi dérive dans la suite. Ce vieillissement en deux étapes sert à la fois à produire les années de revenus qui servent au calcul des prestations et à projeter dans le futur (ou le présent). Plus précisément, à partir de l'ERFS de l'année **N** on construit une pseudo-ERFS pour l'année **N+1**. Pour cette année-là, un ménage **m** de caractéristiques socio-démographiques $X_m^{(N+1)}$ et de revenus $R_m^{(N+1)}$ de la population est représenté dans notre échantillon par un ménage enquêté l'année **N**, de caractéristiques $X_m^{(N)}$ et de revenus $(1+\tau_{N \rightarrow N+1})R_m^{(N)}$ (seconde étape de la dérive), dont la pondération a été modifiée de manière à être représentative de la population en **N+1** (par un calage sur marges qui constitue la première étape de la dérive) et où $\tau_{N \rightarrow N+1}$ est l'évolution moyenne du revenu entre les années **N** et **N+1** (les évolutions dépendent de la nature du revenu¹²). On réitère ensuite l'opération afin de constituer une pseudo-ERFS pour l'année **N+2**.

Les marges de calage qui permettent d'adapter les pondérations et être ainsi représentatif des années **N** et **N+1** sont les suivantes :

- Données démographiques : effectif d'individus par croisement sexe et tranche d'âge, sur le champ des ménages ordinaires au 15 novembre (17 catégories) ;
- Données de logement : nombre de ménages locataires ;

11 Modulo des abattements pour changement de situation (professionnelle ou familiale), les prestations dépendant des revenus de l'antépénultième année sont : les aides au logement, le complément familial, les allocations sous condition de ressource de la PAJE, l'allocation de rentrée scolaire, l'allocation de l'adulte handicapé.

12 En particulier, le modèle fait vieillir les revenus salariaux d'un individu selon l'évolution du salaire mensuel de base de l'ensemble des individus appartenant au même secteur d'activité et à la même catégorie socioprofessionnelle. Ces coefficients sont issus des enquêtes Acemo.

- Données d'emploi : nombre de ménages par type de composition familiale (8 catégories), statut d'activité (5 catégories), catégorie socio-professionnelle (3 catégories) et nombre de ménages comprenant au moins un salarié, un indépendant ou un retraité, tout cela provenant des résultats des enquêtes Emploi ;
- Données fiscales : nombre de déclarations comprenant des salaires, des retraites ou revenus d'indépendants (5 catégories), redressé pour correspondre à un champ France métropolitaine, ménages ordinaires (uniquement pour **N** pour des raisons de disponibilité).

Les revenus sont ensuite revalorisés pour la projection de l'année **N** à l'année **N+1** de manière à ce que :

- l'évolution totale des traitements et salaires, des retraites et des revenus des indépendants correspond à l'évolution constatée des masses déclarées correspondantes (source fiscale),
- l'hétérogénéité de l'évolution des salaires et des revenus de chômage et de préretraite est décrite par croisement secteur-catégorie socio-professionnelle¹³, selon les résultats de l'enquête trimestrielle Acemo,
- les revenus accessoires, des capitaux mobiliers, et les revenus et déficits fonciers correspondent à l'évolution observée des montants par déclaration (source fiscale),
- les plus-values évoluent comme la moyenne annuelle à l'ouverture de l'indice CAC-all-tradable (anciennement SBF250).

En revanche, certaines données ne sont pas disponibles pour la projection de l'année **N+1** à l'année **N+2**. La deuxième étape de la dérive s'effectue alors de la façon suivante :

- les salaires sont revalorisés par croisement secteur-catégorie socio-professionnelle selon les résultats de l'enquête trimestrielle Acemo,
- les allocations chômage suivent l'évolution de l'allocation d'aide au retour à l'emploi (revalorisation de la partie fixe et de la partie minimale),
- les retraites sont revalorisées selon le taux de revalorisation des pensions civiles et militaires,
- les coefficients appliqués aux revenus des indépendants, aux capitaux mobiliers, aux revenus accessoires et aux revenus fonciers correspondent à une moyenne sur les quatre dernières années (évolution tendancielle), à l'exception des revenus agricoles pour lesquels on dispose d'une évolution produite par l'Agreste,
- comme précédemment, les plus-values évoluent comme la moyenne annuelle à l'ouverture de l'indice CAC-all-tradable (anciennement SBF250).

Il est utile de souligner que du fait du vieillissement statique de l'ERFS sur deux années, il n'y a pas de variations brutales de la situation individuelle. Ainsi, le recalage sur marges correspond à l'idée que les situations des chômeurs enquêtés l'année **N** correspondent à celles vécues par les chômeurs en **N+2**, une fois prise en compte l'évolution moyenne des revenus. Cette absence de dynamique individuelle réelle implique un meilleur ciblage des prestations sous conditions de ressources et de l'impôt sur le revenu, puisque l'ordre dans l'échelle des revenus est le même en **N+2** qu'en **N** et que les plus démunis en **N** sont toujours les plus démunis en **N+2**, et *in fine* les mieux ciblés. Dans la réalité, une personne ayant perdu son emploi, paiera l'année suivante un impôt sur le revenu élevé au regard de ses revenus actuels. À l'inverse, une personne dont les revenus ont fortement augmenté sur les deux dernières années pourra tout de même percevoir une aide au logement, malgré ses revenus actuels élevés. Ces situations de décalage dues à la législation sociale et fiscale française n'existent pas dans le modèle Ines du fait du vieillissement statique.

13 On repère alors la catégorie socio-professionnelle et le secteur du dernier emploi occupé pour les chômeurs et préretraités.

Dans une période de retournement conjoncturel, ces transitions individuelles sont plus fréquentes et, ainsi, l'erreur de ciblage que fait le modèle Ines risque d'être plus forte. C'est une difficulté à prendre en compte dans un exercice de *nowcasting* où l'on s'intéresse à des évolutions.

3.2. Détourner Ines pour faire du *nowcasting*

L'objectif de cette étude est de prévoir le taux de pauvreté (ou plutôt son évolution) avec un an d'avance par rapport à l'ERFS, en utilisant la micro-simulation. La méthode envisagée consiste à projeter Ines sur une seule année. La législation de l'année **N+1** est entièrement codée de façon stable à l'été de l'année **N+2**. Par ailleurs, à cette date, la plupart des paramètres de vieillissement (excepté pour les revenus des capitaux mobiliers, les revenus fonciers, et les revenus accessoires) nécessaires au calcul du niveau de vie de **N+1** sont disponibles. Il est donc possible d'estimer des évolutions du taux de pauvreté entre **N** et **N+1** dès que le rebasage d'Ines sur l'ERFS **N** est réalisé. Comme l'ERFS **N** (échantillon noyau) est disponible vers septembre **N+2**, en comptant 1 mois de travail de changement de base, il serait donc possible de gagner 11 mois sur les délais actuels de production de cet indicateur. Si cette méthode est utilisée dès cette année, l'évolution précoce du taux de pauvreté entre 2013 et 2014 pourrait être connue en octobre 2015 au lieu de septembre 2016 (voir illustration 3). Ce calendrier correspond au régime de croisière où la chaîne de production de l'indicateur est bien stabilisée et où aucune difficulté particulière n'est rencontrée lors du changement de millésime de l'ERFS.

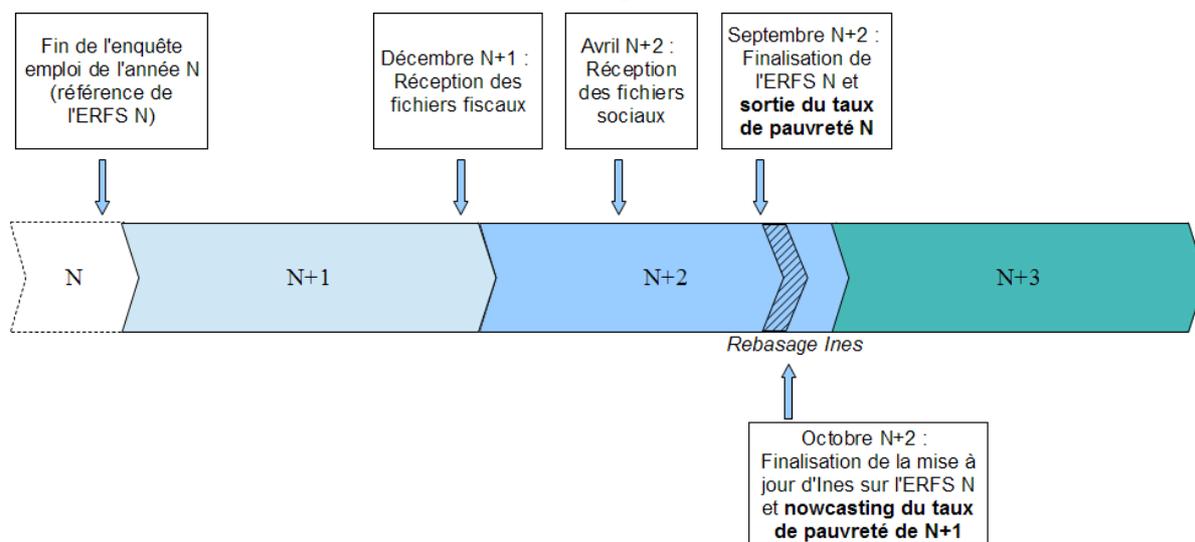


Illustration 3: Déroulement du processus de production de l'ERFS (en haut) et du *nowcasting* du taux de pauvreté N+1 (en bas).

Pour évaluer la qualité de la méthode, nous comparons les évolutions estimées pour la période 2009-2012 à celles produites par les ERFS¹⁴. Or la comparaison entre la microsimulation et la méthode d'appariement n'est pas immédiate et les écarts sont de plusieurs natures. Tout d'abord, il existe des différences conceptuelles sur le calcul du niveau de vie. Ensuite, Ines simule la législation des prestations et des prélèvements sur barème, en imputant éventuellement du non-recours de manière aléatoire, là où l'ERFS utilise des informations appariées réellement observées et en corrigeant la non-réponse par imputation. Enfin, Ines doit projeter la structure socio-démographique et les revenus sur une ou deux années sur la base d'un certain nombre d'hypothèses, alors que l'ERFS utilise la seule information disponible.

14 On utilisera dans la suite les millésimes 2009 à 2011 des ERFS, auxquels on appliquera les législations 2010 à 2012. Le rebasage d'Ines sur l'ERFS 2012 n'ayant pas été terminé à ce jour, il n'est pas encore possible de calculer l'évolution du taux de pauvreté entre 2012 et 2013.

Tout l'enjeu de l'analyse qui suit est de comparer les résultats d'Ines et de l'ERFS à l'aune des indicateurs usuels de niveau de vie et d'inégalités, mais aussi d'expliquer ces différences en faisant la part entre chacun de ces différents facteurs. La méthodologie est la suivante :

1. identifier et neutraliser les différences conceptuelles entrant dans le calcul du niveau de vie entre l'ERFS et Ines (partie 3.4) ;
2. isoler l'effet « micro-simulation » : comparer, à partir d'une même ERFS, les résultats de la simulation sur barème avec les réelles observations de l'ERFS. Pour Ines, on appelle dans la suite cette étape « évaluation contemporaine », car l'année de législation simulée est identique à celle de l'ERFS utilisée (partie 4.1). Afin de simuler les prélèvements et prestations de l'année **N**, ce sont donc des « pseudo-ERFS » **N-1** et **N-2** qui sont produites à partir de l'ERFS de l'année **N**. L'opération n'est donc plus un vieillissement mais un rajeunissement statique sur deux ans. Nous désignons cette démarche par le terme de « dérive inverse » par opposition à la « dérive classique ».
3. isoler l'effet « vieillissement » : comparer pour une même année **N**, les résultats obtenus avec l'ERFS de l'année **N**, et ceux obtenus avec le modèle Ines utilisant l'ERFS de l'année **N-1**. Cette deuxième évaluation est appelée « évaluation décalée » dans la suite, car l'année de législation simulée est décalée d'un an par rapport à celle de l'ERFS utilisée (partie 4.3). La population et les revenus ne seront projetés que sur une seule année au lieu de deux dans l'exercice habituel d'Ines. Par différence avec l'étape précédente, on déduit l'effet « vieillissement ». Pour l'évaluation décalée, c'est l'ERFS de l'année **N-1** qui sert de référence. On construit donc les pseudo-ERFS des années **N-2** d'une part (dérive inverse), et **N** d'autre part (dérive classique).

Afin de réaliser ces deux évaluations à partir d'Ines, le schéma de vieillissement statique classique est ainsi détourné (illustration 4).

Plus précisément, la méthode de projection **N** à **N-1** et **N-1** à **N-2** est de nature identique à celle utilisée pour projeter de **N** à **N+1** (cf partie 3.1) : les marges de calages sont les mêmes et les coefficients de revalorisation des revenus sont recalculés à partir des évolutions utilisées pour projeter de **N** à **N+1** ($\tau_{N \rightarrow N-1} = (1 + \tau_{N-1 \rightarrow N})^{-1} - 1$) et utilisés de manière à respecter les mêmes règles sur l'évolution totale des revenus.

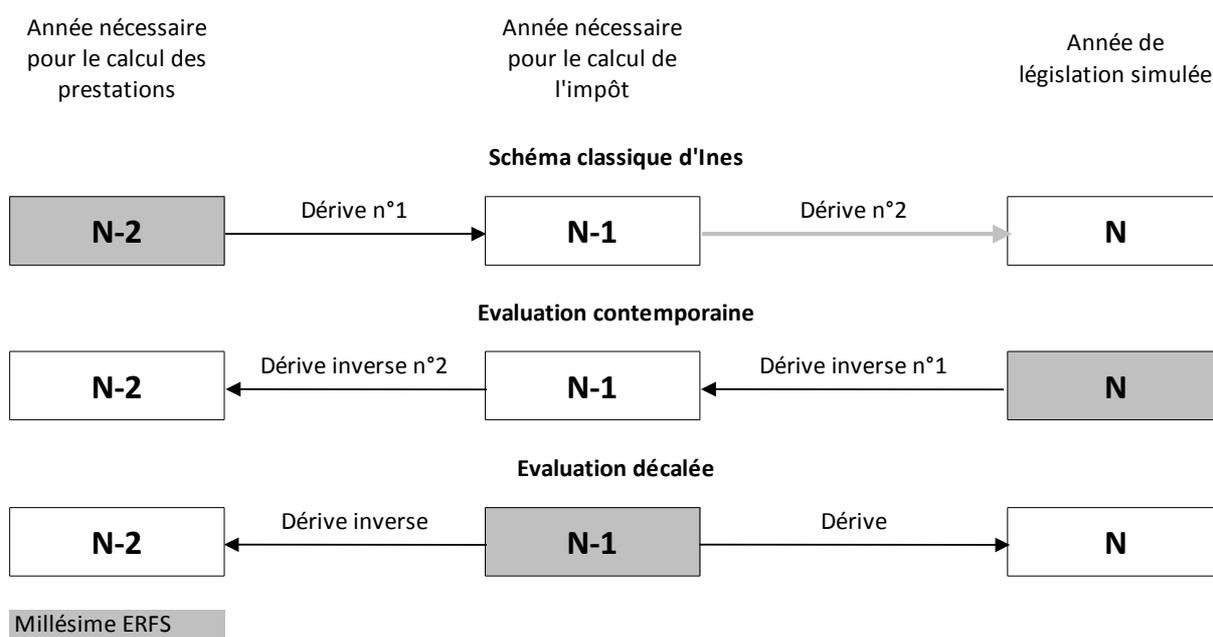


Illustration 4: Détournement de l'utilisation classique d'Ines dans le cadre de l'exercice de nowcasting.

3.3. Définition des estimateurs

L'exercice de *nowcasting*, s'il devenait pérenne, consisterait à réaliser une projection à partir de la dernière information disponible. Son objectif serait de prédire l'évolution $\widehat{\Delta}_{ERFS}$ du taux de pauvreté entre les années **N** et **N+1**, évolution qui sera disponible en **N+3**. Soit $\widehat{\tau}_{(N)}^{ERFS}$ le taux de pauvreté de l'année **N** issu de l'ERFS, l'évolution du taux de pauvreté entre N et N+1 s'écrit :

$$\widehat{\Delta}_{ERFS} = \widehat{\tau}_{(N+1)}^{ERFS} - \widehat{\tau}_{(N)}^{ERFS}$$

Plus précisément, en septembre de l'année **N+2**, l'ERFS de l'année **N** devient disponible. Le rebasage d'Ines qui s'ensuit rend alors possible à la fois l'évaluation *contemporaine* du taux de pauvreté de l'année **N** $\widehat{\tau}_{(N)}^{(N)}$, et l'évaluation décalée du taux de l'année **N+1** $\widehat{\tau}_{(N+1)}^{(N)}$.

De la comparaison entre l'ERFS **N** et l'évaluation *contemporaine* en **N** ($\widehat{\tau}_{(N)}^{ERFS} - \widehat{\tau}_{(N)}^{(N)}$), nous déduisons l'écart lié à la différence de méthode (appariement contre imputation sur barème).

De la comparaison entre l'évaluation contemporaine en **N**, et l'évaluation décalée en **N** qui avait été faite sur la base de l'ERFS **N-1** ($\widehat{\tau}_{(N)}^{(N)} - \widehat{\tau}_{(N)}^{(N-1)}$), nous déduisons l'écart lié au fait d'utiliser une ERFS trop vieille d'un an (approximation liée au vieillissement projetant d'un an les données et variabilité due au changement d'échantillon).

Le *nowcasting* du taux de pauvreté sera réalisé en évolution de façon à éviter les écarts entre les niveaux projetés et observés, la microsimulation ayant tendance à produire un niveau de pauvreté plus faible que les résultats d'appariement.

Au moment où l'exercice de *nowcasting* est réalisé, nous disposons des indicateurs suivants :

- $\widehat{\tau}_{(N)}^{ERFS}$, l'estimation du taux de pauvreté de N calculé à partir de l'ERFS N ,
- $\widehat{\tau}_{(N+1)}^{(N)}$, l'estimateur de l'évaluation décalée de la législation $N+1$ à partir de l'ERFS N ,
- $\widehat{\tau}_{(N)}^{(N)}$, l'estimateur de l'évaluation contemporaine de la législation N à partir de l'ERFS N ,
- $\widehat{\tau}_{(N)}^{(N-1)}$, l'estimateur de l'évaluation décalée de la législation N à partir de l'ERFS $N-1$.

Pour estimer $\widehat{\Delta}_{ERFS}$, trois estimateurs sont envisageables.

1. Le premier est $\widehat{\tau}_{(N+1)}^{(N)} - \widehat{\tau}_{(N)}^{ERFS}$, c'est-à-dire la différence entre l'évaluation décalée pour $N+1$ et le taux de l'ERFS pour N . Nous écartons d'emblée cette solution qui est entachée de deux types d'erreurs : l'erreur due au fait de comparer des estimateurs obtenus par des méthodologies très différentes (nous verrons à la section 4.2.2 que la micro-simulation produit des taux de pauvreté inférieurs d'environ 0,5 point à ceux obtenus par l'ERFS¹⁵), et l'erreur due à la méthodologie de vieillissement transformant une population représentative de l'année N en une pseudo-population $N+1$ (de l'ordre de 0,2 point¹⁶).
2. Le deuxième est $\widehat{\tau}_{(N+1)}^{(N)} - \widehat{\tau}_{(N)}^{(N-1)}$, à savoir la différence entre deux évaluations décalées. Cet estimateur a l'avantage d'avoir strictement la même méthodologie entre les deux termes de la différence (cf illustration 4) : le fait de raisonner en différence permet donc de neutraliser l'écart quasi-systématique entre la méthode de micro-simulation et l'appariement direct de l'ERFS. En revanche, cet indicateur comporte l'inconvénient d'utiliser deux ERFS différentes, ce qui introduit une variabilité en espérance nulle, mais potentiellement importante dans l'indicateur différentiel produit.
3. Le troisième estimateur de $\widehat{\Delta}_{ERFS}$ envisageable est $\widehat{\tau}_{(N+1)}^{(N)} - \widehat{\tau}_{(N)}^{(N)}$, à savoir l'écart entre l'évaluation *décalée* en $N+1$ et l'évaluation contemporaine en N . Comme pour le deuxième estimateur, celui-ci comporte l'avantage de neutraliser les écarts liés à la différence de méthodologie. En revanche, c'est bien la même ERFS (celle de l'année N) qui est utilisée pour les deux termes de la différence, ce qui permet de contrôler la variabilité qui serait purement liée au changement d'échantillon entre deux ERFS.

Dans la suite, c'est le dernier estimateur qui est choisi : $\widehat{\Delta}_{nowcasting} = \widehat{\tau}_{(N+1)}^{(N)} - \widehat{\tau}_{(N)}^{(N)}$. En effet, nous estimons que la différence de méthodologie de vieillissement entre l'évaluation contemporaine et l'évaluation décalée est de second ordre par rapport à l'imprécision liée au fait d'utiliser deux ERFS différentes. Bien que potentiellement légèrement biaisé, le troisième estimateur nous semble plus précis¹⁷.

15 Confer Tableau 6, écarts entre les colonnes (1) et (2)

16 Confer Tableau 6, écarts entre les colonnes (2) et (3)

17 Il faut garder par la suite à l'esprit que l'estimation de l'évolution du taux de pauvreté produite par différence entre deux ERFS est entachée de la variabilité due au changement d'échantillon, de l'ordre de 0,6 point.

3.4. Différences conceptuelles de niveau de vie entre la microsimulation et l'appariement

Comme mentionné précédemment à propos de l'impôt sur le revenu, il existe quelques différences de champ entre Ines et l'ERFS. Certaines de ces différences peuvent être contrôlées par des modifications *ad hoc* dans Ines. Ainsi, lorsque cela est possible et à condition que la logique générale du modèle ne soit pas bouleversée, Ines est légèrement modifié pour l'exercice, afin de se rapprocher au maximum de la méthodologie de l'ERFS. Parfois, ces modifications conduisent à dégrader le niveau de précision du modèle.

- La première différence concerne l'**échantillon source**. Depuis le passage à l'enquête Emploi en continu (en 2003), Ines s'appuie sur une version élargie de l'ERFS. En effet, l'ERFS dans sa version « noyau » comprend dans son champ les ménages interrogés au 4^e trimestre de l'année **N**. Dans sa version élargie, la taille de l'échantillon double presque en incluant tous les ménages en dernière interrogation pendant l'un des trois premiers trimestres de l'année **N** (ménages sortants) et symétriquement les ménages en première interrogation sur l'un des trois premiers trimestres de l'année **N+1** (ménages entrants). Dans cet exercice de *nowcasting*, le modèle Ines a été adapté pour ne traiter que les ménages du noyau.
- Dans l'ERFS, aucun traitement n'est effectué pour corriger le fait que l'on ne retrouve pas forcément toutes les déclarations fiscales lorsqu'un mariage, un pacs ou un divorce a eu lieu dans l'année (ce qui est le cas jusqu'en 2012 avec la suppression de la triple déclaration). En revanche dans le modèle de micro-simulation, la **déclaration manquante** est simulée par simple règle de proportionnalité par type de revenu. Ainsi, la masse totale de revenu est légèrement supérieure dans le modèle Ines que dans l'ERFS. Cette différence n'est pas corrigée puisqu'elle entraînerait une adaptation trop importante du modèle.
- L'**impôt sur le revenu** considéré dans le calcul du revenu disponible n'est pas le même dans l'ERFS et dans Ines : dans l'ERFS il s'agit de l'impôt payé en **N+1** sur les revenus de l'année étudiée (**N**) alors que dans Ines il s'agit de l'impôt payé en **N** sur les revenus de l'année précédente. Une adaptation est effectuée pour calculer l'impôt sur la même année que l'ERFS¹⁸.
- La **prime pour l'emploi** (PPE) est bien, dans les deux cas, celle versée en **N** en fonction des revenus de l'année **N-1**. Normalement, la PPE est résiduelle du montant de RSA activité perçu en **N-1**. Cependant ce calcul n'est pas effectué dans les ERFS (cette correction est prévue pour la production de l'ERFS 2013). Pour se rapprocher de cette méthode, bien qu'il s'agisse d'une approximation, on adapte le modèle Ines en ne retirant pas non plus le montant de RSA du montant de PPE simulé dans Ines. En 2013, 5 % des bénéficiaires de la PPE avaient reçu un montant positif de RSA activité l'année précédente.
- Dans l'ERFS, le niveau de vie tient compte des **revenus du patrimoine** perçus l'année considérée et des prélèvements sociaux qui leur sont associés, exactement comme si tous ces revenus étaient prélevés à la source. Ce calcul s'éloigne de la législation réelle, puisque certains de ces revenus sont effectivement bien imposés à la source (plus-values de cession et valeurs et capitaux mobiliers), alors que d'autres ne sont imposés que l'année suivante (gains de levée d'option), les uns et les autres étant déclarés séparément. Dans Ines (en flux courant) sont pris en compte dans le niveau de vie de l'année **N**, à la fois les prélèvements payés en **N** sur une fraction de ces revenus perçus en **N-1**, et les prélèvements payés en **N** sur une fraction de ces revenus perçus en **N** et

18 À terme, ERFS va adopter la même méthode que dans Ines.

prélevés à la source. Pour l'exercice de *nowcasting*, on s'écarte ici de la législation et du fonctionnement habituel d'Ines afin de se rapprocher de la méthodologie adoptée pour l'ERFS. Les deux méthodes sont donc similaires sur ce point.

- Contrairement à l'ERFS, certains **transferts en nature** sont calculés dans Ines (bourses du collège et lycée, complément de mode de garde de la Paje, Apa, CMU-C¹⁹, ACS). Cependant, ces dispositifs ne sont pas pris en compte dans le calcul du revenu disponible pour Ines.
- Dans les ERFS antérieures à 2012, les **pensions alimentaires versées** incluses dans le niveau de vie sont comptabilisées dans un agrégat fiscal (PSA) qui tient compte de plafonnements réalisés pour le calcul de la réduction d'impôt. Il s'agit d'une erreur corrigée pour l'ERFS 2013. Dans Ines, c'est bien le montant déclaré (avant plafonnement) qui est considéré. Aucune adaptation n'est effectuée. De même, il existe une erreur dans le calcul des revenus accessoires dans l'ERFS, qui n'est pas corrigée.
- Enfin, le **concept d'âge** diffère selon les deux sources. Dans Ines, on calcule l'âge au 31 décembre de l'année, alors que dans l'ERFS, il s'agit de l'âge au moment de l'enquête. Cela peut modifier les résultats sur le calcul des unités de consommation, mais probablement dans une faible mesure puisqu'il s'agit du quatrième trimestre de l'année (le nombre d'unités de consommation est plus grand dans Ines).

19 Dans Ines, seule l'éligibilité à la CMU-C est calculée actuellement. Ce transfert en nature n'est actuellement pas monétarisé.

4. Résultats

Avant de présenter les résultats du *nowcasting*, il convient de comparer les résultats observés de l'ERFS **N** et ceux de l'évaluation *contemporaine* en **N**, afin de faire apparaître l'effet des imputations sur barème.

4.1. Comparaison entre les résultats de l'ERFS et l'évaluation contemporaine

Dans cette partie, nous comparons les résultats de l'ERFS d'une année donnée avec ceux produits par l'évaluation contemporaine (les taux de pauvreté produits par l'une et l'autre de ces méthodes sont donc respectivement $\widehat{\tau}_{(N)}^{ERFS}$ et $\widehat{\tau}_{(N)}^{(N)}$). Puisque ces deux estimateurs se fondent sur la même ERFS, la différence entre les deux correspond au fait d'imputer sur barème plutôt que d'utiliser la méthode d'appariement.

4.1.1. Estimation du niveau de vie et de ses composantes

À partir d'une même ERFS, le fait de simuler les transferts à partir des législations plutôt que de prendre directement l'information appariée conduit à une estimation proche du niveau de vie médian (annexe 5), de +175 à +300 euros par an selon l'année considérée (2009 à 2011), soit une surestimation d'au plus 2 %. Les seuils de pauvreté à 50 % et à 60 % sont, par définition, surestimés dans les mêmes proportions.

Parmi les composantes du niveau de vie (tableau 4), les deux méthodes donnent des résultats assez proches pour un certain nombre de prestations familiales et pour les allocations logement. En revanche, les montants de minima sociaux (en particulier de RSA), de prime pour l'emploi, d'impôt sur le revenu et de prélèvements sur les revenus du patrimoine, sont plus élevés dans Ines que dans l'ERFS. Plus précisément, nous relevons ici les écarts les plus significatifs :

- **RSA socle** : le nombre annuel de bénéficiaires du RSA socle est beaucoup plus élevé dans Ines que dans l'ERFS (+30 à +40 % de bénéficiaires selon l'année considérée²⁰). Le défaut de représentation des allocataires du RSA est structurel dans l'ERFS, et ce malgré une imputation réalisée selon une méthode donneur/receveur (*hot deck*). Pour l'imputation faite dans l'ERFS, la population receveuse est constituée des individus qui déclarent percevoir le RSA dans l'enquête Emploi sur un des trimestres d'interrogation. Les donneurs, eux, sont les allocataires du RSA socle qui ont été retrouvés dans les fichiers Caf ou MSA. Cependant, cette imputation n'augmente le nombre d'allocataires que de 10 % environ, ce qui semble insuffisant puisque même après *hot deck*, le nombre d'allocataires du RSA socle dans l'ERFS est inférieur d'environ 25 à 30 % (selon l'année) à une cible Cnaf rendue comparable²¹. Cette sous-représentation peut être due à une sous-déclaration dans l'Enquête Emploi de la perception de minima sociaux, et éventuellement à la non-réponse de bénéficiaires de minima sociaux à l'enquête Emploi. Le modèle Ines, quant à lui, détermine l'éligibilité au RSA en fonction des ressources du foyer, et n'impute pas de non-recours au dispositif, excepté pour les personnes éligibles à la fois au RSA socle et au RSA activité²². Ce faisant, Ines estime un nombre annuel de

20 Le RSA étant imputé de manière trimestrielle dans Ines, il est préférable de comparer les effectifs trimestriels d'Ines aux cibles de la Cnaf pour juger de la qualité de la méthode. Dans l'ERFS en revanche, c'est l'effectif annuel de bénéficiaires qu'il faut comparer à une cible annuelle Cnaf, le détail des trimestres n'étant pas disponible.

21 Un coefficient *ad hoc* est appliqué pour se rapprocher du champ des ménages ordinaires de France métropolitaine. Par ailleurs, l'effectif annuel est estimé à partir de l'effectif de bénéficiaires au quatrième trimestre, à partir d'un taux de turn-over issu de l'Echantillon National des Allocataires (Ena).

22 Le modèle Ines impute un non-recours au RSA activité en distinguant deux sous-populations parmi ses éligibles : ceux qui sont éligibles au RSA activité seul, et ceux qui sont éligibles à la fois au RSA socle et au RSA activité. Le tirage des recourants est ensuite effectué de manière à coller à une cible externe de

bénéficiaires du RSA socle beaucoup plus proche de l'effectif-cible (entre -10 % et -5 % environ selon l'année). L'écart de montant annuel total de RSA socle entre l'ERFS et Ines est moins accentué que l'écart de bénéficiaires²³ (de l'ordre de +10 % dans Ines).

- **RSA activité** : le montant total de RSA activité est respectivement plus élevé dans Ines par rapport à l'ERFS de +19 % et +9 % pour les années 2010 et 2011. Là encore, l'imputation par *hot deck* réalisée pour l'ERFS augmente faiblement (de 5 % environ) le nombre annuel de bénéficiaires, qui cette fois se trouve relativement proche d'une cible « Cnaf corrigée » (-10 % à +5 % selon l'année). Les bénéficiaires de ce dispositif constituent une population en activité, moins marginalisée que les bénéficiaires du RSA socle, et leur éventuelle sous-représentation dans l'échantillon de l'enquête Emploi serait moins problématique. Le modèle Ines, lui, une fois après avoir déterminé l'éligibilité au RSA activité, impute un non-recours à une partie des éligibles, de manière à se rapprocher d'effectifs-cibles de bénéficiaires pour chaque trimestre. Si le nombre trimestriel de bénéficiaires représente donc bien la réalité, l'effectif annuel, lui, est sous-estimé avec la micro-simulation, là encore en raison d'une prise en compte trop imparfaite des trajectoires infra-annuelles.
- **Impôt sur le revenu (avant PPE)** : une fois neutralisé le décalage de l'année d'imposition dans le modèle Ines, les résultats en termes de masses totales d'impôt sur le revenu prélevées sont proches entre les deux méthodes : +0 % à +8 % selon l'année considérée. Pour toutes les années, la microsimulation donne un nombre de ménages imposés (avant PPE et prélèvement forfaitaire) très légèrement inférieur à celui de l'ERFS (entre -1 et -3 %).
- **Prélèvements sur les revenus du patrimoine** : Ines produit une estimation de ces prélèvements légèrement plus élevée que l'ERFS, de +11 à +16 % sur les masses prélevées et de +6 à +7 % sur le nombre de foyers concernés. Cet écart est potentiellement lié au fait que le modèle de microsimulation tient compte de l'imposition sur les plus-values (prélèvements sociaux et impôt sur le revenu), ce qui n'est jusqu'à présent pas le cas dans l'ERFS.

bénéficiaires par trimestre. En revanche, aucun non-recours au RSA socle n'est imputé pour les éligibles au RSA socle seul. Au total, dans Ines, 10 à 15 % des individus repérés comme étant éligibles au RSA socle n'y recourent pas (uniquement des individus étant aussi éligibles au RSA activité).

23 Une des raisons possibles à cela serait un lissage trop important des revenus avec la méthodologie de trimestrialisation des ressources effectuée dans Ines, ce qui se traduirait par une éligibilité au RSA socle plus fréquente, mais avec des montants individuels plus faibles. Cette explication est vraisemblable car le taux de turnover du RSA socle est plus faible dans Ines que dans l'Ena (27 % contre 42 % en 2011). Les variations individuelles de revenus au cours de l'année, qui sont utilisées pour la trimestrialisation des ressources, seraient donc sous-estimées dans Ines.

	Cible Externe	ERFS 2011	Évaluation contemporaine 2011	Écart Ines / ERFS	Écart ERFS/Cible
<i>En milliards d'euros</i>	(1)	(2)	(3)	[(3)-(2)]/(2)	[(2)-(1)]/(1)
Prélèvements	-	110,1	114,8	4%	-
Impôt sur le revenu (net de crédits d'impôt mais brut de PPE)	50,2	54,9	57,9	5%	9%
Prélèvement forfaitaire	-	3,0	3,2	7%	-
PPE résiduelle	2,8	2,7	3,1	15%	-4%
CSG	-	22,5	22,6	0%	-
Taxe d'habitation	18,7	14,5	14,6	1%	-22%
CRDS	-	4,9	5,1	4%	-
Prélèvements sociaux sur les revenus du patrimoine	-	13,0	14,5	12%	-
Prestations	-	51,3	51,2	0%	-
Allocations familiales	11,8	11,7	11,7	0%	-1%
Complément familial	1,6	1,5	1,5	0%	-6%
Allocations logement	15,2	14,1	14,5	3%	-7%
AAH et son complément	6,4	6,2	5,7	-8%	-3%
RSA socle	6,6	4,9	5,5	12%	-26%
RSA activité	1,4	1,3	1,2	-8%	-7%
Aspa/Asi	1,9	1,7	1,6	-6%	-11%
Paje base et prime de naissance	4,7	4,7	4,8	2%	0%
Allocation de rentrée scolaire	1,4	1,4	1,3	-7%	0%
AEEH	-	0,6	0,6	0%	-
ASF	1,1	1,0	1,2	20%	-9%
CLCA	2,1	2,2	1,6	-27%	5%

Tableau 4 : Montant total des prélèvements et prestations en 2011 calculés dans Ines et ERFS en 2011 et cibles externes.

Champ : France métropolitaine, ménages ordinaires.

Source des cibles externes : Cnaf pour les prestations, annuaire statistique des finances publiques pour les prélèvements.

4.1.2. Indicateurs de pauvreté et d'inégalités

Si les niveaux de vie médians (et donc les seuils de pauvreté) sont proches entre la méthode d'appariement et la micro-simulation quelle que soit l'année considérée, il n'est pas évident qu'il en soit de même pour les indicateurs de distribution, tels que le taux de pauvreté. En effet, la part de personnes vivant sous le seuil de pauvreté n'évolue en principe pas systématiquement comme ledit seuil.

Entre 2009 et 2011, l'évaluation contemporaine conduit à produire un taux de pauvreté à 60 % inférieur à celui obtenu d'après l'ERFS par la méthode d'appariement, avec un écart de -0,3 à -0,6 point de pourcentage selon l'année considérée (cf illustration 5). Cet écart est donc supérieur à l'intervalle de confiance du taux de pauvreté à 60 % fourni par l'ERFS (+/- 0,3 point). Le taux de pauvreté à 50 % est, lui, assez fortement sous-estimé, avec un écart de -0,4 à -0,8 point de pourcentage (cf Tableau 7)²⁴.

Cet écart de niveau peut s'expliquer en partie par les écarts sur les bénéficiaires de RSA socle (Ines ayant 28 % de bénéficiaires de plus que l'ERFS en 2011). En effet, lorsque l'on retire aléatoirement (tirage aléatoire simple) le bénéfice du RSA socle à 28 % des ménages d'Ines en 2011, le taux de pauvreté à 60 % augmente de 0,1 point et celui à 50 % augmente de 0,3 point (atteignant respectivement 14,0 % et 7,5 % en 2011). Selon la répartition de cet écart en fonction du niveau de revenu, il est possible que cela affecte plus ou moins un taux ou l'autre. Dans le cas

²⁴ Notons qu'en 2011, une modification a eu lieu sur l'ERFS concernant l'imputation des revenus des produits financiers. En effet, depuis l'ERFS 2005, ces revenus étaient imputés à partir de l'enquête Patrimoine 2004, en quatre étapes (modèle de détention des différents types de produits financiers, simulation des montants pour les ménages détenteurs d'après l'ERFS, application de taux de rendement moyens et recalage). À partir de l'ERFS 2011, la méthode d'imputation reste globalement inchangée, mais s'appuie sur les données de l'enquête Patrimoine 2010. Un double point est alors fourni de manière rétroactive pour l'ERFS 2010 (annexe 7). Ce rebasage a, à lui seul, un léger effet à la baisse sur le niveau de vie médian (-120 euros), et à la hausse sur le taux de pauvreté (+0,16 point).

de ce tirage aléatoire simple, l'intensité de la pauvreté passe de 82,6 % à 82,0 % dans Ines, à comparer à 81,1 % dans l'ERFS. En effet, une proportion importante d'allocataires du RSA socle est au-dessus du seuil de pauvreté à 60 % (39 % des allocataires en 2012 d'après l'ERFS). Une part importante d'entre eux passerait en dessous du seuil de pauvreté s'ils ne recouraient pas au RSA (ni socle ni activité alors). On estime qu'environ 150 000 ménages seraient dans cette situation en 2012 d'après l'ERFS, soit 0,6 % de la population totale. L'examen de cas-types (avec les barèmes de 2014) tendent à dire qu'il s'agirait pour beaucoup de familles monoparentales ou de couples mono-actifs avec au moins deux enfants.

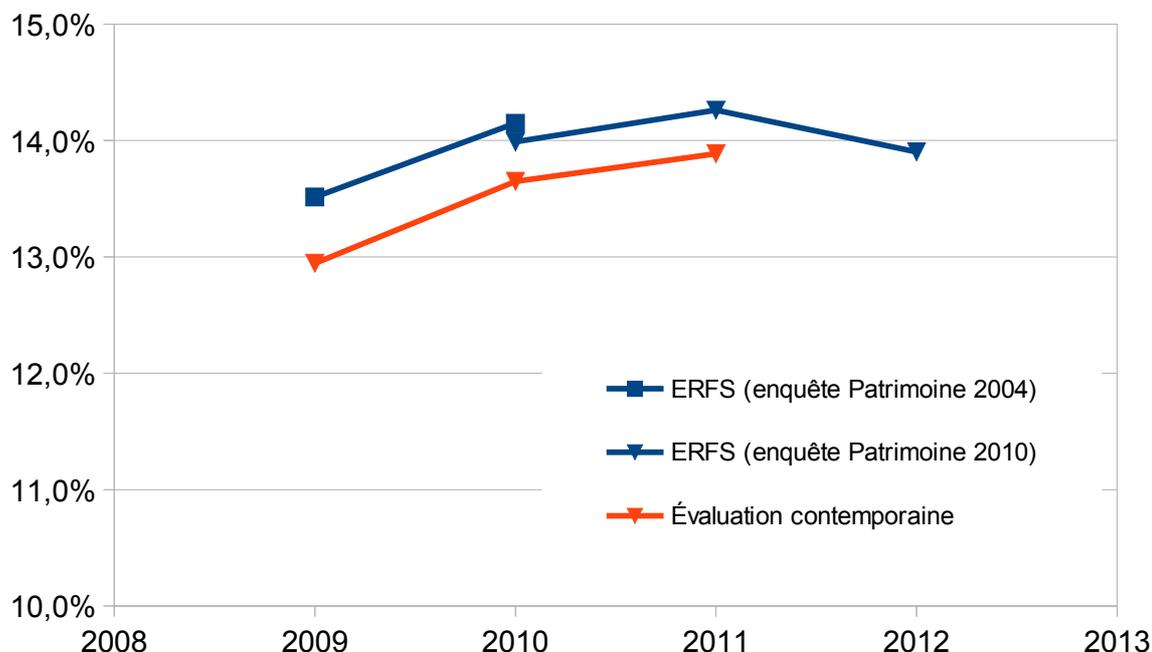


Illustration 5 : Taux de pauvreté à 60 % estimés par l'ERFS et par Ines en évaluation contemporaine

Champ : ménages ordinaires de France métropolitaine

Les courbes de taux de pauvreté à 50 % ou à 60 % issues de l'une ou l'autre des deux méthodes sont relativement parallèles. Autrement dit, les évolutions des taux de pauvreté d'une année à l'autre sont globalement bien prédites par la micro-simulation, l'écart observé sur le taux de pauvreté s'expliquant davantage par des différences structurelles (plus de bénéficiaires du RSA socle ou du RMI avant 2009 dans Ines que dans l'ERFS, notamment).

L'intensité de la pauvreté, en revanche, est légèrement surestimée par la micro-simulation par rapport à la méthode d'appariement, quelle que soit l'année considérée (de l'ordre de +1,5 points de pourcentage entre 2009 et 2011, soit environ +2 %).

Fait rassurant, la distribution des niveaux de vie est très peu dépendante de la méthode, même si sur le cinquième percentile de niveau de vie, l'écart atteint 4 % (annexe 6).

4.2. Comparaison entre les résultats de l'ERFS et ceux de l'évaluation décalée

Dans cette partie, nous comparons les résultats de l'évaluation décalée du taux de pauvreté en $N+1$ produite à partir de l'ERFS d'une année N , avec ceux directement obtenus avec l'ERFS de l'année $N+1$, disponibles un an plus tard (les taux de pauvreté produits par l'une et l'autre de ces méthodes sont donc respectivement $\widehat{\tau}_{(N+1)}^{ERFS}$ et $\widehat{\tau}_{(N+1)}^{(N)}$). Puisque ces deux estimateurs ne se

fondent pas sur la même ERFS, la différence entre les deux correspond d'une part au fait d'imputer sur barème plutôt que d'utiliser la méthode d'appariement (cf. partie 4.1), mais aussi d'autre part à l'imperfection du vieillissement des données.

4.2.1. Estimation du niveau de vie et de ses composantes

Avec l'évaluation décalée d'Ines qui s'appuie sur l'ERFS qui précède l'année dont on veut mesurer le taux de pauvreté, on obtient un niveau de vie médian (annexe 5) supérieur à celui de l'ERFS disponible un an plus tard. L'écart varie entre +190 à +580 euros par unité de consommation selon l'année considérée (2010 à 2012), soit au maximum une surestimation de 3 %. Les seuils de pauvreté à 50 % et à 60 % sont, par définition, surestimés dans les mêmes proportions. Par rapport à l'évaluation contemporaine, le fait d'utiliser une ERFS plus vieille d'un an induit une révision de 0 à +250 euros par unité de consommation. Cet effet sur le niveau de vie médian et sur les seuils de pauvreté est en moyenne plus faible que celui dû au fait d'utiliser la micro-simulation plutôt que l'appariement.

Parmi les composantes du niveau de vie (annexe 6), les écarts exposés plus haut ne sont pas bouleversés : les montants de minima sociaux (exceptés d'allocation adulte handicapé, AAH), de prime pour l'emploi, d'impôt sur le revenu et de prélèvements sur les revenus du patrimoine, sont plus élevés dans Ines que dans l'ERFS. Les prestations familiales, les allocations logement, et les prélèvements forfaitaires, peuvent être sur ou sous-estimés par rapport à l'ERFS, selon l'année. De manière générale, l'écart introduit en utilisant une ERFS différente est de second ordre sur les niveaux par rapport à l'écart introduit en utilisant la micro-simulation plutôt que la méthode d'appariement, même si pour certains dispositifs et certaines années, les résultats peuvent déroger à cette généralité. Plus précisément, on peut dresser le bilan suivant pour les dispositifs qui marquent le plus d'écarts :

- **RSA socle** : environ +25 % de bénéficiaires dans l'évaluation décalée par rapport à l'ERFS, soit un écart légèrement réduit (de l'ordre de -5%) par rapport à celui exposé en première étape ;
- **RSA activité, impôt sur le revenu (avant PPE), PPE** : les écarts avec l'ERFS sont très similaires à ceux décrits en première étape, le fait d'utiliser une autre ERFS plus vieille influençant peu les résultats grâce à la méthode de vieillissement ;
- **AAH** : pour ce minimum social, l'écart avec l'ERFS est amplifié par le fait d'utiliser une ERFS plus ancienne pour l'année 2010 (-9 % sur la masse versée dans l'évaluation décalée au lieu de -2 % dans l'évaluation contemporaine). Les résultats s'éloignent également des effectifs cibles de la Cnaf, rendus comparables. Cela serait lié à des évolutions également assez fluctuantes de ce dispositif dans l'ERFS, fluctuations que le vieillissement réalisé pour l'exercice ne suffirait pas à corriger ;
- **Prélèvement forfaitaire libératoire/obligatoire** : avec l'évaluation contemporaine, ce prélèvement était estimé dans les mêmes proportions que l'ERFS pour les années 2010 et 2011 (+4 % à +7 %). Avec l'évaluation décalée, l'écart avec l'ERFS s'inverse et s'établit à -15 % environ pour ces deux années. Pour l'année 2012, l'évaluation décalée produit une masse de prélèvements plus élevée que celle de l'ERFS de près de 50 %. Outre la volatilité de ces revenus que l'on appréhende mal dans le vieillissement, cet écart exceptionnel pourrait être partiellement lié à des changements de taux intervenus en 2012 concernant l'imposition de ces revenus [CAZ14], entraînant un report conséquent du prélèvement forfaitaire vers l'imposition au barème ;
- **Prélèvements sur les revenus du patrimoine** : l'évaluation décalée produit un résultat assez éloigné de l'ERFS (jusqu'à +16 % pour l'année 2010 et +21 % pour l'année 2012), comme c'était déjà le cas avec l'évaluation contemporaine en première étape. Là encore, la volatilité importante de ces revenus d'une année à l'autre pourrait expliquer l'écart généré par le décalage d'ERFS.

En résumé, l'évaluation décalée n'introduit que peu d'écarts à l'ERFS par rapport à l'évaluation contemporaine. Autrement dit, le vieillissement réalisé pour le *nowcasting* est assez satisfaisant puisqu'il parvient à rendre compte assez fidèlement de la situation de l'année suivante. Néanmoins, plus les revenus sont volatiles d'une année à l'autre, plus le fait d'utiliser une information moins fraîche nuira à la qualité des résultats. Dans la même veine, une sous-représentation ponctuelle d'une frange de la population dans l'enquête Emploi pourra avoir des conséquences sur la prévision du taux de pauvreté de l'année suivante, même si ces effets sont en principe atténués par la partie « calage sur marges » du vieillissement d'Ines.

4.2.2. Estimation des indicateurs de pauvreté et d'inégalités

Pour les années 2010 à 2012, les taux de pauvreté en *décalé* sont estimés respectivement à -0,6, -0,5 et -0,3 point de pourcentage par rapport à ceux de l'ERFS (colonnes (3)-(1) du tableau 6). Sur les deux années 2010 et 2011, l'écart lié à la micro-simulation est plus important que l'écart lié au fait d'utiliser une ERFS vieille d'un an.

Le taux de pauvreté à 50 % est, lui, également sous-estimé, de -0,7 à -1,0 point de pourcentage par rapport à l'ERFS (tableau 7). Ici, l'écart semble intégralement lié à la microsimulation, l'évaluation décalée ne dégradant quasiment pas l'estimation.

En termes de niveau encore, l'intensité de la pauvreté à 60 % est bien prédite par l'évaluation décalée, avec un écart à l'ERFS qui ne bouge pas par rapport à la première étape. Enfin, les indicateurs classiques de distribution (indices inter-quantiles et indice de Gini) ne sont pas bouleversés par le fait d'utiliser une ERFS plus vieille d'un an, même si cette opération va plutôt dans le sens de s'éloigner des résultats de l'ERFS (Tableau 5).

	ERFS	Évaluation contemporaine	Évaluation décalée
Q3/Q1	1,88	1,87	1,87
P80/P20	2,21	2,20	2,18
D9/D1	3,56	3,47	3,43
P95/P5	5,53	5,31	5,20
Gini	0,306	0,300	0,295

Tableau 5 : Estimation des principaux indicateurs d'inégalités pour l'année 2011

Champ : ménages ordinaires de France métropolitaine

4.3. Nowcasting avec mise en conditions réelles

Si l'exercice conduit dans cette étude était amené à être renouvelé de façon pérenne, il consisterait à produire la variation du taux de pauvreté entre l'année **N** et l'année **N+1** dès que l'ERFS de l'année **N** est disponible, autrement dit à l'automne de l'année **N+2**. En effet, à cette date, la législation de l'année **N+1** est déjà entièrement codée de façon stable. L'objet de cette partie est donc de se placer dans ces conditions réelles.

Néanmoins, pour les années 2009 à 2012 considérées ici, on dispose en 2015 de certaines informations plus récentes que celle dont on aurait disposé si on avait réellement effectué l'exercice en conditions réelles. En effet, pour l'évaluation d'une année **N**, les coefficients d'évolution des revenus de capitaux mobiliers, des revenus fonciers, et des revenus accessoires, ne sont reçues qu'au printemps de l'année **N+2**, et ne sont donc pas mobilisables à l'automne de l'année **N+1**. Dans le présent document de travail, les dérivées sont tout de même réalisées à partir des meilleures informations disponibles, alors qu'en conditions réelles, nous aurions dû faire des hypothèses d'évolutions tendanciennes.

Comme vu plus haut en partie 3.3, notre estimateur de *nowcasting* correspond à l'écart entre $\tau_{(N+1)}^{(N)}$, l'évaluation décalée de l'année **N+1** à partir de l'ERFS **N**, et $\tau_{(N)}^{(N)}$, l'évaluation contemporaine à partir de l'ERFS **N**.

Le sens de l'évolution annuelle du taux de pauvreté aurait été correctement prédit à chaque fois pour le taux à 60 % (Tableau 6), mais pas toujours pour le taux à 50 % (Tableau 7). En 2012, l'évaluation décalée aurait en effet prédit une baisse du taux de pauvreté à 50 % (-0,1 point), là où l'ERFS fournit une hausse de 0,2 point. En 2012, notre méthode aurait conduit à estimer une diminution de 0,33 point du taux de pauvreté à 60 %, très proche de celle observée avec l'ERFS (-0,36 point). Pour 2012, les évolutions simulées du taux de pauvreté sur les tranches d'âge 18-24 ans et 75 ans et plus sont également satisfaisantes. Pour les années 2010 et 2011, les évolutions prédites vont dans le même sens que celles de l'ERFS²⁵, mais elles sont plus molles. Cela peut traduire le fait que l'absence de dynamique dans le modèle est plus gênante dans les périodes de crise, où les transitions individuelles sont plus fréquentes.

Un autre point de comparaison est donné par le dispositif Filosofi. Le taux de pauvreté à 60 % évolue de +0,1 point entre 2010 et 2011 puis de -0,03 point entre 2011 et 2012 pour atteindre 14,27 points. Le coefficient de Gini quant à lui passe de 0,301 en 2010 à 0,302 en 2011 puis à 0,298 en 2012. Les évolutions prédites par Ines sont donc comprises pour ces deux années entre celles de l'ERFS et celles de Filosofi.

Taux de pauvreté monétaire À 60 %	ERFS	Évaluation contemporaine	Évaluation décalée	Estimateur du <i>nowcasting</i>	Cible observée
N	(1)	(2)	(3)	$(3)_N - (2)_{N-1}$	$(1)_N - (1)_{N-1}$
2009	13,5*	12,9			
2010	14,0 (14,1*)	13,6	13,4	0,4	0,6
2011	14,3	13,9	13,8	0,1	0,3
2012	13,9		13,6	-0,3	-0,4

* : résultats obtenus à partir de l'enquête Patrimoine 2004 (à défaut, enquête Patrimoine 2010)

Tableau 6 : Taux de pauvreté à 60 % estimés par l'ERFS et par Ines entre 2009 et 2012 (en %).
Champ : ménages ordinaires de France métropolitaine.

Note de lecture : Entre 2011 et 2012, l'exercice de *nowcasting* mené sur l'ERFS 2011 aurait conduit à prédire une évolution de -0,3 point du taux de pauvreté à 60 % (de 13,9 % à 13,6%).

Taux de pauvreté monétaire À 50 %	ERFS	Évaluation contemporaine	Évaluation décalée	Estimateur du <i>nowcasting</i>	Cible observée
N	(1)	(2)	(3)	$(3)_N - (2)_{N-1}$	$(1)_N - (1)_{N-1}$
2009	7,5*	6,7			
2010	7,7 (7,8*)	7,0	6,9	0,2	0,4
2011	7,9	7,2	7,3	0,2	0,3
2012	8,1		7,1	-0,1	0,2

* : résultats obtenus à partir de l'enquête Patrimoine 2004 (à défaut, enquête Patrimoine 2010)

Tableau 7 : Taux de pauvreté à 50 % estimés par l'ERFS et par Ines entre 2009 et 2012 (en %).
Champ : ménages ordinaires de France métropolitaine.

Note de lecture : Entre 2011 et 2012, l'exercice de *nowcasting* mené sur l'ERFS 2011 aurait conduit à prédire une évolution de -0,1 point du taux de pauvreté à 50 % (de 7,2 % à 7,1 %).

25 A l'exception du taux de pauvreté des 18-24 ans en 2010.

De même, entre 2010 et 2012, le sens de l'évolution annuelle de l'indice de Gini aurait toujours été correctement prédit par le *nowcasting*, mais la méthode conduit à prévoir des évolutions d'ampleur plus faible (Tableau 8).

Indice de Gini	ERFS	Évaluation contemporaine	Évaluation décalée	Estimateur du <i>nowcasting</i>	Cible observée
N	(1)	(2)	(3)	$(3)_N - (2)_{N-1}$	$(1)_N - (1)_{N-1}$
2009	0,290*	0,282			
2010	0,303 (0,299*)	0,293	0,284	0,002	0,009
2011	0,306	0,300	0,295	0,002	0,003
2012	0,303		0,298	-0,002	-0,003

* : résultats obtenus à partir de l'enquête Patrimoine 2004 (à défaut, enquête Patrimoine 2010)

Tableau 8 : Indices de Gini estimés par l'ERFS et par Ines entre 2009 et 2012 (en %).

Champ : ménages ordinaires de France métropolitaine.

Note de lecture : Entre 2011 et 2012, l'exercice de *nowcasting* mené sur l'ERFS 2011 aurait conduit à prédire une évolution de -0,002 de l'indice de Gini (de 0,300 à 0,298).

Notons que des estimations ont également été produites pour des sous-catégories de populations précises (tranche d'âge des individus, type de composition familiale). Les résultats sont globalement satisfaisants (confer annexes 8, 9 et 10), hormis pour l'année 2011 où des divergences auraient été observées entre l'évolution prédite et l'évolution observée, pour la plupart des types de ménage. Plus précisément, les taux de pauvreté prédits par l'estimateur du *nowcasting* sont moins bons pour les sous-catégories les plus petites. Ainsi, les résultats les plus mauvais sont observés sur les catégories des personnes seules de moins de 65 ans, des couples sans enfants, et des familles monoparentales, chacune de ces catégories représentant moins de 10 % des ménages.

5. Conclusion et enjeux

Le *nowcasting* avec Ines conduirait à donner une estimation du taux de pauvreté 11 mois plus tôt qu'actuellement. La méthode par micro-simulation produit, de 2009 à 2012, un taux de pauvreté à 60 % plus faible en niveau que celui de l'ERFS, de l'ordre de 0,3 à 0,6 point. Cet écart de niveau est lié en partie au fait d'utiliser la micro-simulation plutôt que l'information appariée (avec notamment une meilleure estimation dans Ines du RSA socle, les bénéficiaires se retrouvant moins souvent dans la population pauvre), et pour une autre partie au fait d'utiliser une base de microsimulation vieille d'un an.

Pour les transferts sociaux les plus classiques entrant dans la composition du niveau de vie (impôt sur le revenu, prestations familiales, minima sociaux, allocations logement...), les écarts entre l'évaluation contemporaine et l'évaluation décalée sont plus réduits que ceux entre la méthode d'appariement et l'évaluation contemporaine. Mais pour certains transferts liés à des revenus très conjoncturels (revenus financiers en particulier), la méthodologie de vieillissement d'Ines ne parvient pas à rendre compte d'une volatilité suffisante.

Afin de neutraliser ces écarts de niveau structurels, il est plus pertinent de raisonner sur des évolutions. Sur la période de 2010 à 2012, le *nowcasting* aurait toujours conduit à prédire des évolutions du taux de pauvreté à 60 % dans le même sens qu'observées un an plus tard avec l'ERFS. Néanmoins, les évolutions prédites sont toujours légèrement plus faibles (en valeur absolue) que celles de référence. Cette mollesse de nos indicateurs peut traduire le problème de l'absence de dynamique dans le modèle Ines, plus gênante en temps de retournement conjoncturel.

Les résultats de l'exercice sont également concluants sur les taux de pauvreté à 60 % par tranche d'âge (notamment chez les moins de 25 ans et chez les plus de 65 ans). Ils sont relativement concluants aussi par type de ménage, même si les prévisions se détériorent lorsque l'on s'intéresse à des sous-populations plus petites. L'évolution de l'indice de Gini aurait également toujours été prédite dans le bon sens. En revanche, le *nowcast* des évolutions des autres indicateurs d'inégalités est moins satisfaisant (taux de pauvreté à 50 % et intensité de la pauvreté à 60%). Cette moindre qualité des estimations peut s'expliquer par l'écart plus important entre Ines et l'ERFS au bas de la distribution des niveaux de vie, sans doute en raison de la divergence de représentation des bénéficiaires du RSA socle.

La prudence est de rigueur dans l'interprétation d'un tel indicateur d'évolution. Tout d'abord, si l'exercice venait à se pérenniser, la vigilance devra être accrue sur les revenus et les dispositifs les plus volatiles d'une année à l'autre, et ce d'autant plus que pour certains de ces revenus (revenus des capitaux mobiliers en particulier), la dérive ne pourra être effectuée que sur la base d'une hypothèse d'évolution tendancielle. Les estimations d'évolutions des indicateurs d'inégalités seront plus spécifiquement tributaires de cette hypothèse, dans la mesure où les revenus les plus variables d'une année à l'autre étant majoritairement détenus par les individus les plus aisés. Les estimations d'évolutions des indicateurs de pauvreté, elles, y seront moins sensibles. Ensuite, les estimations délivrées seront à considérer avec précaution en période de crise ou de reprise marquées, pendant lesquelles le vieillissement effectué sur une ERFS datant d'un an ne suffit pas forcément à traduire la réalité. Enfin, le lecteur doit garder à l'esprit la sensibilité du taux de pauvreté aux méthodes de simulation des prestations sociales, méthodes susceptibles d'être modifiées d'une année sur l'autre dans un modèle Ines en perpétuelle évolution. Les indicateurs délivrés ne seraient donc en aucun cas comparables l'année sur l'autre.

Pour conclure, dans l'hypothèse de pérennisation d'un tel estimateur, la réflexion déjà engagée sur le rapprochement méthodologique entre Ines et l'ERFS (abandon de l'utilisation de l'échantillon élargi pour Ines, modification du millésime d'impôt sur le revenu pour l'ERFS) devra se poursuivre.

6. Bibliographie

- [HOU14] Houdré C., Ponceau J., Zergat Bonnin M., Les niveaux de vie en 2012, Insee Première, 2014.
- [BOZ12] Bozio A., Fabre B., Goupille J., Lafféter Q., Le modèle de micro-simulation TAXIPP – Version 0.2, Guide Méthodologique IPP, 2012.
- [NAV13] Navicke J., Rastrigina O., Sutherland H., Nowcasting Indicators of Poverty Risk in the European Union: A Microsimulation Approach, Euromod Working paper No. EM 11/13, 2013.
- [LEV13] Leventi C., Navicke J., Rastrigina O., Sutherland, H., Nowcasting: Estimating developments as regards the risk of poverty and income distribution in 2012 and 2013, Research note 1/2013, 2013.
- [AVR13] Avram S., Bouvard L., Euromod Country Report: France 2009-2013, Euromod country report, 2013.
- [CAZ14] Cazenave M.-C., Duval J., Fontaine M., Stehlé J., Redistribution : en 2013, les nouvelles mesures accroissent la fiscalité des ménages et réduisent légèrement les inégalités, France, Portrait Social, 2014.
- [LeM04] Le Minez, S., Une estimation des taux de pauvreté monétaire avec le modèle de microsimulation INES adossé aux enquêtes Revenus fiscaux, Note Drees, 2004.
- [BUF10] Buffeteau S., Cazenave M.-C., Duval J., Projection des taux de pauvreté pour le rapport remis au parlement, Note Drees, 2010.
- [EID12] Eidelman, A., Le modèle Ines : tests de validité, comparaison avec l'ERFS et réflexions sur l'avenir, Note Insee n° 1747/DG75-F120/, 2012.
- [MIR08] Mirouse, B., Evolution des indicateurs de pauvreté d'ici à 2011, suivant différentes hypothèses, Note Drees, 2008.

Annexes

Annexe 1 Premiers travaux de *nowcasting* menés par la Drees en 2004

Les premiers travaux s'apparentant à un exercice de *nowcasting* par micro-simulation des niveaux de vie au niveau français remontent à 2004 [LeM04]. Ils ont été réalisés avec le modèle de micro-simulation Ines 2002. À cette époque, le modèle reposait sur l'ancêtre de l'ERFS, l'enquête Revenus Fiscaux (ERF) de 1999. L'ERF était le résultat de l'appariement entre les revenus de l'année 1999 et les résultats de l'enquête Emploi de mars 2000, alors de cadence annuelle. À l'époque, le modèle Ines fonctionnait de la manière suivante (illustration 6) :

1. projection sur deux ans de l'ERF 1999 grâce à deux calages sur marges successifs pour obtenir une « pseudo-ERF 2001 », c'est-à-dire une base de données de contenu analogue à celle de l'ERF en 2001 ;
2. évolution individuelle des revenus entre 1999 et 2001 selon la catégorie des revenus en appliquant des coefficients multiplicateurs idoines ;
3. simulation de l'impôt sur le revenu payé en 2002 sur les revenus de 2001 et des prestations sociales perçues en 2002, également sur la base des revenus de 2001 ;
4. multiplication des revenus de 2001 par l'inflation entre 2001 et 2002 pour en déduire un revenu de 2002.

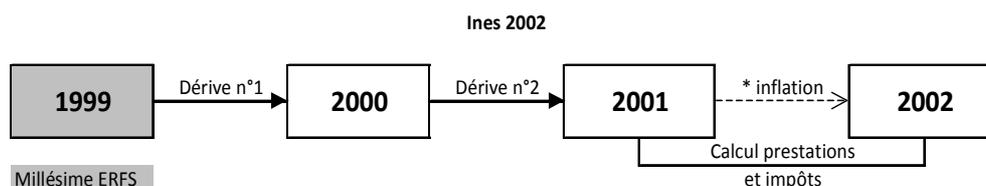


Illustration 6: Fonctionnement d'Ines en 2002.

On peut se demander à quelle quantité il est plus pertinent de comparer le niveau de vie calculé en bout de chaîne à partir des revenus projetés de 2001, de l'inflation entre 2001 et 2002, et des prestations et prélèvements respectivement reçus et payés en 2002

De prime abord, le niveau de vie ainsi décrit semble conceptuellement comparable à un niveau de vie de 2002. Toutefois, la simple multiplication des revenus bruts de 2001 par l'inflation ne permet pas de tenir compte des déformations de la distribution des revenus ni de la variété des situations individuelles. Même si le niveau des prix de 2002 est pris en compte, le niveau de vie obtenu est conceptuellement plus proche de celui de 2001, à la différence près que les prestations et prélèvements sont calculés à partir des barèmes de 2002. De plus, dans l'ERF 2001, l'impôt considéré dans le calcul du niveau de vie de 2001 est celui qui est payé en 2002 sur les revenus de 2001²⁶ (convention de « provisionnement » de l'impôt sur le revenu), alors que dans Ines, cet impôt intervient dans le calcul du niveau de vie de 2002 (en adoptant une convention de flux courant). Au total, la seule différence conceptuelle résiduelle entre Ines 2002 et les résultats de l'ERF 2001 en euros 2002 résiderait dans l'année de prise en compte des prestations.

Les seuils et taux de pauvreté des ERF 2001 et 2002 n'étaient pas présentés dans la note (non encore produits au moment de sa rédaction), mais il est possible aujourd'hui de comparer les résultats avec les ERF. Le seuil de pauvreté pour 2002 estimé avec le modèle de micro-simulation est supérieur à la fois aux seuils produits par l'ERF pour 2001 (+3,4 %) et 2002

²⁶ À ce jour, l'ERFS d'une année *N* prend toujours en compte dans le revenu disponible de l'année *N* l'impôt payé en *N+1* relatif aux revenus de l'année *N*.

(+1,1 %). Comme ce seuil est simplement proportionnel au niveau de vie médian, ce dernier est estimé à un niveau plus élevé pour le modèle de micro-simulation.

	ERF 2001	ERF 2002	Ines 2002
Seuil de pauvreté	613	627	634
Taux de pauvreté	6,1%	6,0%	6,6%

Tableau 9: Seuil (en euros 2002 par mois) et taux de pauvreté à 50 % du niveau de vie médian.

Champ: France métropolitaine, ménages ordinaires.

Comme vu plus haut, cet écart de 3,4 % entre les niveaux de vie d'Ines 2002 et d'ERF 2001 s'explique en partie par l'inflation (1,9 % en 2002), et en partie par le fait que les prestations sont plus élevées dans Ines 2002 que dans ERF 2001 (car calculées avec le barème de 2002). Cependant, ces deux éléments sont loin d'expliquer l'intégralité de l'écart. De nombreux facteurs peuvent également contribuer à expliquer cette différence et les résultats présentés dans la note sont trop frustes pour conclure. Il est possible que les revenus aient été augmentés trop fortement lors de l'étape de vieillissement de l'échantillon avec le modèle de micro-simulation. Il est également possible que les prélèvements soient trop faibles dans Ines²⁷, et/ou que les prestations soient trop élevées²⁸.

Le taux de pauvreté à 50 % est, quant à lui, également supérieur lorsqu'il est produit avec le modèle de micro-simulation (+0,5 point par rapport à l'ERF 2001). Cet écart est supérieur à l'intervalle de significativité de l'ERF (le taux de pauvreté étant donné à plus ou moins 0,3 point), et ne peut s'expliquer uniquement par des différences d'année des prestations sociales à prendre en compte dans le calcul du niveau de vie. En effet, comme expliqué ci-dessus, les masses de prestations versées devraient être supérieures à celles de l'ERF 2001. Partant du principe que les prestations sont plus ciblées vers le bas de la distribution, le taux de pauvreté obtenu par micro-simulation devrait donner une population moins pauvre que celle de l'ERF, ce qui n'est pas le cas.

27 La sous-estimation des prélèvements serait d'autant plus forte qu'à la différence de l'ERF, Ines inclut la taxe d'habitation.

28 Là aussi, Ines, à cette époque, se distinguait de l'ERF, en incluant dans le revenu disponible les aides à la garde de jeunes enfants et en excluant les allocations logement pour les accédants à la propriété. Compte tenu des masses relatives de ces prestations et de leur progressivité, la redistributivité des prestations sociales devrait en principe être supérieure dans le modèle de micro-simulation.

Annexe 2

Travaux de *forecasting* de la Drees en 2008

En juin 2008²⁹, un chiffrage de la Drees a été produit pour le cabinet de Martin Hirsch, haut commissaire aux solidarités actives contre la pauvreté [MIR08]. Il présente les résultats d'un exercice de *forecasting*, pour donner l'évolution des indicateurs de pauvreté entre 2006 et 2011 sous différents scénarios macroéconomiques, avec ou sans réforme des minima sociaux (revenu de solidarité active (RSA) et minimum vieillesse).

La méthodologie est la suivante : le modèle de micro-simulation est utilisé tout d'abord dans sa version habituelle, c'est-à-dire en simulant au niveau de chaque ménage les niveaux de vie de 2006, en s'appuyant sur l'ERFS 2004. Les conditions d'emploi de 2011 selon 3 scénarios différents sont ensuite utilisées pour projeter la structure de l'emploi 2011 sur l'échantillon représentatif de 2006, avec création d'emplois et augmentation moyenne des salaires réels. Ces scénarios étaient les suivants :

- le scénario bas (au fil de l'eau) du pacte de stabilité et de croissance (augmentation de 7,7 % de tous les salaires et création de 360 000 emplois) ;
- le scénario haut du pacte de stabilité et de croissance (augmentation de 7,7 % de tous les salaires et création de 755 000 emplois) ;
- un scénario tendanciel qui reprend les évolutions constatées sur les 5 dernières années (augmentation de 4,6 % de tous les salaires et création de 686 000 emplois) ;.

Le reste des caractéristiques de la population est inchangé. Les seules modifications du système social et fiscal concernent la mise en place du RSA et la revalorisation du minimum vieillesse.

	Taux de pauvreté relatif à 60 % du niveau de vie médian		
	Ensemble	< 18 ans	> 65 ans
Projection INES pour 2006	11,2%	15,1%	9,2%
ERFS 2006 (observé)	13,2%	NC	NC
Scénario tendanciel pour 2011	9,2%	12,3%	9,1%
Scénario bas pour 2011	9,8%	12,9%	9,9%
Scénario haut pour 2011	9,5%	12,4%	10,3%
ERFS 2011 (observé)	14,3%	19,5%	NC

Tableau 10: Résultats du forecasting de 2011 à partir d'Ines 2006 (ie ERFS 2004) dans les différents scénarios économiques, avec mise en place du RSA et revalorisation du minimum vieillesse.

Champ : France métropolitaine.

Source : Ines 2006 et ERFS 2006 et 2011.

Le taux de pauvreté à 60 % estimé par le modèle de micro-simulation (*nowcasting* sur 2 ans à partir de l'ERFS 2004) atteint 11,2 % dans Ines alors qu'il a été évalué à 13,2 % avec l'ERFS 2006, soit un écart important. Selon le scénario pris en compte, le taux de pauvreté projeté en 2011, avec mise en place du RSA et augmentation du minimum vieillesse varie entre 9,2 % et 9,8 % (contre 14,3 % dans l'ERFS 2011).

Là encore, les informations sont trop frustes pour conclure sur la qualité de l'exercice. Les hypothèses retenues d'évolution des salaires réels et de création d'emploi se sont révélées par la suite trop optimistes. En particulier, malgré la croissance démographique, l'emploi total n'a pas sensiblement augmenté sur la période (26,1 millions de personnes en emploi en 2006 et 26,4 millions en 2011). Les salaires nets mensuels moyens des postes à temps complet n'ont évolué que de 3,3 % entre 2006 et 2010. De cet exercice, il faut retenir la forte sensibilité des résultats aux hypothèses retenues et l'impossibilité de prévoir à horizon lointain l'évolution des dispositifs de prélèvements et prestations sociales.

29 À ce moment de l'année, le rebasage du modèle Ines sur l'ERFS 2005 n'était pas encore finalisé (cf illustration 1).

Annexe 3 Projections des taux de pauvreté par la Drees en 2009

En 2009, la Drees a produit des projections d'évolution des taux de pauvreté entre 2007 et 2009 pour le 2^e rapport remis au Parlement relatif au suivi de l'objectif de baisse d'un tiers de la pauvreté en cinq ans. Une note rédigée en 2010 précise les modalités de l'exercice [BUF10]. Il s'agissait d'utiliser le modèle Ines de manière conventionnelle pour produire des niveaux de vie de 2008 à partir de l'ERFS 2006. Ensuite, l'année 2009 était simulée sans calage et avec les modifications législatives connues (introduction du RSA activité mais pas de prime de solidarité active) et avec un scénario de destruction d'emploi (451 000 emplois détruits), en ciblant les salariés les plus fragiles sur le marché du travail (les emplois détruits concernaient pour 15 % des emplois intérimaires, et pour 23 % des contrats à durée déterminée).

	ERFS (observé)		Ines 2009 (simulé sur ERFS 2006)	
	2007/2008	2008/2009	2007/2008	2008/2009
Taux de pauvreté relatif à 60 %	-0,4 pt	+0,5 pt	-0,4 pt	-0,7 pt
Intensité de la pauvreté relative	+0,3 pt	+0,5 pt	-0,1 pt	-0,8 pt
Taux de pauvreté ancrée en 2006	-0,9 pt	+0,2 pt	-0,3 pt	-1,6 pt

Tableau 11: Evolution d'indicateurs de pauvreté obtenu avec les ERFS 2007, 2008 et 2009 (à gauche) et avec le modèle de microsimulation Ines à partir de l'ERFS 2006 (à droite) dans son utilisation usuelle pour 2007 et 2008 et avec un scénario de destruction d'emploi en 2009.

Champ : France métropolitaine.

Source : Ines 2009 (à partir de l'ERFS 2006) et ERFS 2007 à 2009.

Seules les évolutions des taux sont présentées et non les niveaux, pour éviter d'afficher les écarts de niveau avec ceux des ERFS. Les évolutions simulées entre 2007 et 2008 sont assez satisfaisantes. Le taux de pauvreté relatif à 60 % diminue de la même ampleur entre les ERFS 2007 et 2008 (-0,4 point). Le taux de pauvreté ancrée en 2006 diminue plus vivement entre les deux ERFS que par micro-simulation. L'intensité de la pauvreté relative présente des évolutions divergentes.

Les projections réalisées en 2009 pour 2009 ne donnent pas des résultats convaincants sur l'évolution de ces indicateurs. En effet, le *nowcasting* par micro-simulation (avec un scénario de destruction d'emplois) prévoyait une diminution de 0,7 point du taux de pauvreté relatif à 60 %, aux antipodes de l'augmentation constatée de 0,5 point entre les ERFS 2008 et 2009. Les évolutions prédites des autres indicateurs ne sont pas plus convaincantes.

Nous pouvons en conclure que le calage sur marges et la projection ont donné une évolution du taux de pauvreté satisfaisante entre 2007 et 2008. En revanche, le scénario de destructions d'emplois, pourtant raisonnable (451 000 emplois détruits contre 336 000 d'après l'Insee dans les Comptes de la nation 2010), ne permet pas d'aboutir à des évolutions satisfaisantes des indicateurs projetés, même qualitativement.

Annexe 4 Comparaison entre appariement et micro-simulation par l'Insee en 2012

En août 2012, une note de bilan sur le modèle Ines et ses limites a été rédigée [EID12]. Une des parties de ce document traite de la comparaison des résultats de calculs sur barème des prélèvements et prestations sociales (modèle Ines) aux résultats issus de l'appariement (données de prestations sociales et de prélèvements de l'ERFS). L'analyse portait sur l'année 2009.

L'objectif de l'exercice n'était pas de réaliser une projection du taux de pauvreté, mais seulement de mesurer l'écart à la réalité introduit par la micro-simulation, en faisant l'effort de neutraliser certains gros facteurs potentiels d'écarts. Ainsi, cette fois le champ d'étude est restreint aux mêmes prestations sociales que dans l'ERFS (les transferts en nature tels que l'allocation personnalisée d'autonomie, la subvention crèche et les aides à la garde, bien que modélisées dans Ines, ne sont pas pris en compte). Les calculs sont faits uniquement sur l'échantillon « noyau » de l'ERFS (c'est-à-dire l'appariement fait avec les ménages de l'enquête emploi au 4^e trimestre). Cependant, certaines sources de différences ne sont pas neutralisées. Par exemple, l'impôt sur le revenu pris en compte n'est pas celui de la même année. Par ailleurs, en 2012, le traitement des valeurs mobilières différait entre ERFS et Ines, ce qui a été modifié depuis.

Dans ce document, le taux de pauvreté à 60 % du niveau de vie médian obtenu par le modèle de micro-simulation vaut 12,7 % alors qu'il vaut 13,5 % par appariement des sources fiscales et sociales³⁰. Les niveaux de vie moyens sont comparables, celui obtenu par micro-simulation étant inférieur de seulement 26 euros par UC par an.

Plus en détail :

- l'impôt sur le revenu (hors crédits d'impôt) de l'ERFS est supérieur de 2,1 milliards d'euros (soit 4,6 %) à celui d'Ines. Cet écart est réduit de moitié lorsque l'on prend les crédits d'impôt en compte (hors PPE). Il ne s'agit toutefois pas des mêmes années puisque l'impôt sur le revenu de l'ERFS est celui acquitté en 2010 sur les revenus 2009, alors que dans Ines il s'agit de l'impôt sur le revenu acquitté en 2009 sur les revenus 2008 ;
- un gros écart porte sur la prime pour l'emploi (PPE) : l'ERFS verse 1 milliard de PPE de moins que cette version d'Ines, alors qu'il s'agit bien de la même année ; pourtant, dans l'ERFS, la PPE n'est pas nette du RSA activité perçu l'année précédente, ce qui devrait rendre son montant plus important que celui d'Ines ;
- l'ERFS décrit 1,7 milliard d'euros de prélèvements libératoires, contre 4,9 milliards dans cette version d'Ines. Cet écart était vraisemblablement dû au fait que le modèle calculait un prélèvement sur les revenus du patrimoine imputés³¹ alors que l'ERFS prenait en

30 Le niveau du taux de pauvreté est un indicateur à prendre avec précaution, car sensible aux petites variations du seuil de pauvreté. Ainsi, avec un seuil de pauvreté qui serait défini à 61 % du niveau de vie médian au lieu de 60 %, on passerait d'un taux de pauvreté de 12,7 % à 13,6 % avec la microsimulation (le saut est du même ordre avec l'ERFS).

31 En effet, comme l'ERFS s'appuie sur les déclarations fiscales pour calculer les revenus disponibles, les revenus non fiscalisés ne sont pas appréhendés par la source. Il s'agit en particulier des revenus générés par les livrets jeune, d'épargne populaire, des livrets A et ex-CODEVI, des comptes épargne logement, des plans d'épargne logement, des plans épargne action (exonérés sous certaines conditions). Lors de la production de l'ERFS, les revenus de ces produits financiers sont imputés. De même, les revenus annuels de contrats de capitalisation et d'assurance-vie sont imputés, puisqu'ils ne sont imposés qu'au dénouement (selon le type de dénouement, la durée de détention, le type du contrat) et donc connus qu'avec retard et accumulés sur plusieurs années. En revanche, l'imposition de ces revenus, de caractère exceptionnel, n'est pas soustraite dans l'ERFS de l'impôt total (contrairement au traitement qui est fait des plus-values) et ainsi le niveau de vie contient un revenu lissé de contrats d'assurance-vie mais une contrepartie en termes d'imposition qui ne l'est pas. À cette époque, pour être cohérent dans le traitement lissé de ces revenus, le modèle Ines corrigeait l'imposition en imputant un prélèvement moyen sur les revenus lissés de contrats d'assurance-vie. Ce faisant, et ne tenant compte ni de la durée de détention, ni du type de contrat, ni du type de dénouement à venir, ni des modifications à venir du régime d'imposition de ces contrats, tous inconnus, le prélèvement était

compte les prélèvements mentionnés dans les déclarations fiscales. Le modèle Ines s'est depuis adapté à ce formalisme ;

- le modèle Ines attribue 6,9 % de plus de prestations sociales que l'ERFS (3,6 milliards d'euros), se répartissant en 1,7 milliard d'allocation logement de plus, 700 millions de plus d'allocations familiales, 600 millions de plus de RSA socle (la prime de solidarité active n'était alors pas prise en compte dans le modèle Ines), ou encore 400 millions de plus de complément familial. La majorité des prestations ont des masses plus importantes dans Ines, à l'exception de l'allocation d'éducation de l'enfant handicapé (AEEH) ou de l'allocation de soutien familial (ASF).

	Cible Externe	ERFS 2009	Ines barèmes 2009 sur ERFS 2009	Écart Ines / ERFS	Écart ERFS/Cible
<i>En milliards d'euros</i>					
	(1)	(2)	(3)	[(3)-(2)]/(2)	[(2)-(1)]/(1)
Prélèvements	-	137,1	142,5	4%	-
Impôt sur le revenu	-	44,4	45,5	3%	-
Prélèvements libératoires	-	1,7	4,9	194%	-
PPE résiduelle	-	3,4	4,4	29%	-
CSG	-	65,3	65,2	0%	-
Taxe d'habitation	-	13,3	13,3	0%	-
CRDS	-	4,6	4,8	4%	-
Prélèvements sociaux sur les revenus du patrimoine	-	11,3	13,2	17%	-
Prestations	-	48,6	52,2	7%	-
Allocations familiales	11,9	11,2	11,9	6%	-6%
Complément familial	1,6	1,4	1,8	26%	-12%
Allocations logement	14,6	13,2	14,9	13%	-10%
AAH et son complément	5,5	5,4	5,7	5%	-3%
RMI, API, RSA socle, PSA et prime de décembre	6,0	5,2	5,8	12%	-12%
RSA activité à partir du 1 ^{er} juin 2009	0,6	0,5	0,7	27%	-14%
Aspa/Asi	-	1,7	1,8	5%	-
Paje base et prime de naissance	4,7	4,6	4,6	0%	-1%
All. de rentrée scolaire et prime exc. de 2009	1,8	1,7	1,8	5%	-8%
AEEH	-	0,6	0,3	-46%	-
ASF	1,1	0,9	0,8	-12%	-16%
CLCA	2,2	2,2	2,2	0%	0%

Tableau 12: Montant total des transferts en 2009 en euros dans la comparaison Ines/ERFS réalisée en 2012

Champ : ménages ordinaires, de France métropolitaine, dont la personne de référence n'est pas étudiante et dont le revenu disponible n'est pas négatif [EID12].

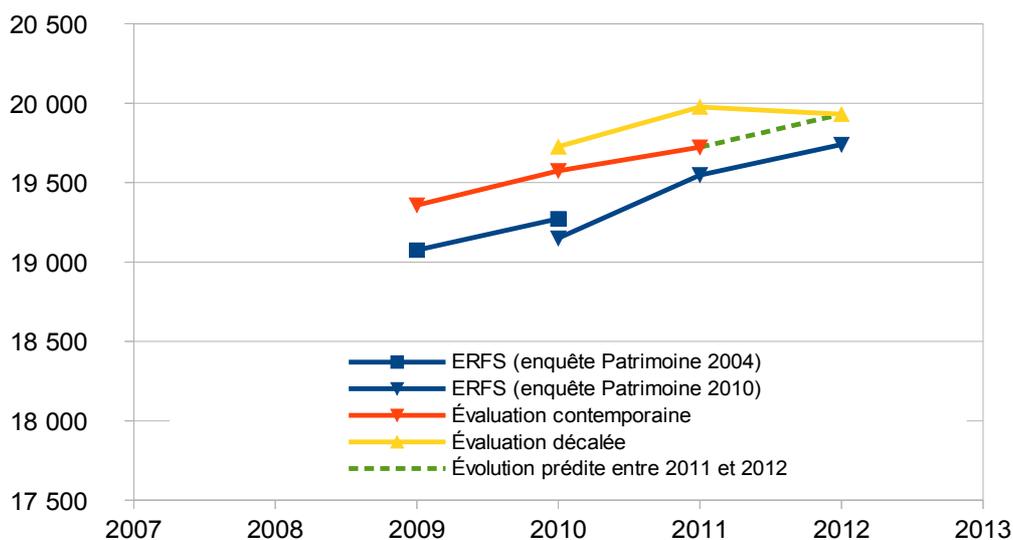
Note : Les cibles externes sont estimées par les soins des auteurs sur un champ France métropolitaine, ménages ordinaires, donc pas tout à fait comparables. Dans cette version du modèle Ines, la prime de solidarité active et la prime exceptionnelle d'allocation de rentrée scolaire n'étaient pas simulées.

En conclusion, certains écarts entre Ines et ERFS sont très probablement imputables à des différences méthodologiques : millésime d'impôt à inclure dans le revenu disponible, soustraction du RSA activité pour obtenir la prime pour l'emploi résiduelle, prise en compte de l'imposition des revenus d'assurance-vie, de la prime de solidarité active, du complément exceptionnel à l'allocation de rentrée scolaire. D'autres résultats font en revanche apparaître des différences plus structurelles et irréconciliables : éligibilité à l'allocation de soutien familial, à l'allocation d'éducation de l'enfant handicapé. Il est en effet très difficile d'identifier le niveau de handicap dans l'enquête emploi et dans les sources fiscales, d'où la difficulté de micro-simuler les dispositifs liés au handicap. De même, il est impossible de déduire la situation en termes de

fixé à un niveau assez arbitraire. Par ailleurs, la montée en charge des contrats d'assurance-vie sur les dernières années provoquait un décalage entre l'impôt réellement perçu une année donnée sur les revenus accumulés sur la durée de détention et le prélèvement hypothétique obtenu par l'imposition au même taux des revenus annuels des encours actuels. Le modèle Ines a depuis été modifié pour traiter les revenus des contrats d'assurance-vie de la même manière que l'ERFS.

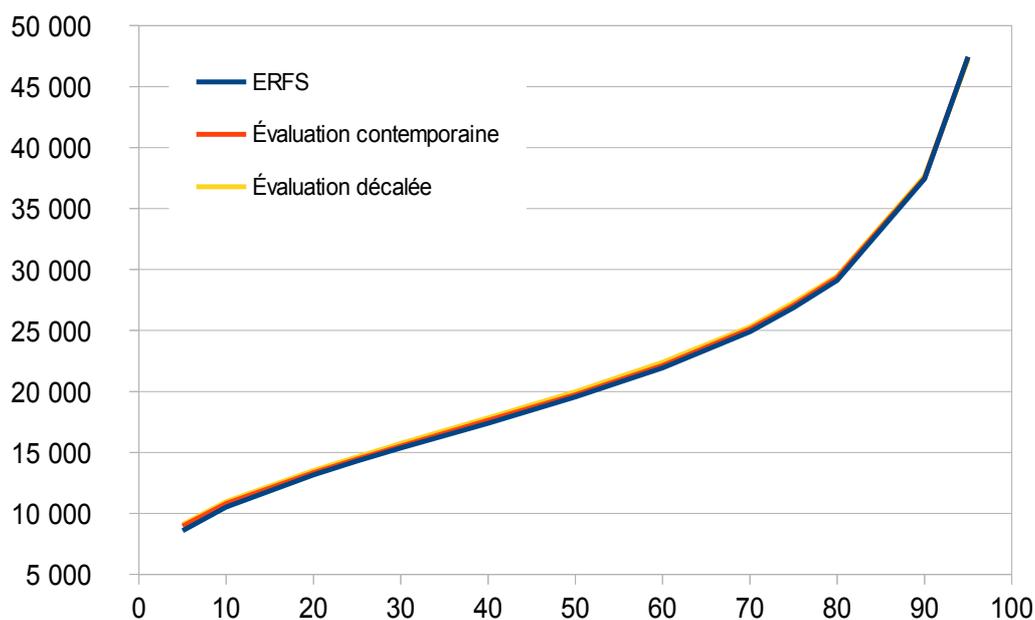
pension alimentaire due ou de reconnaissance de l'enfant élevé par un parent isolé, ce qui pose problème pour le calcul de l'ASF. Enfin, on peut également identifier dès à présent que l'ERFS ne propose pas forcément des montants de prestations sociales plus convaincants. En effet, il existe un déficit important de masse concernant le RSA et d'autres allocations.

Annexe 5 Niveaux de vie médian (euros courants) d'après les trois méthodes envisagées



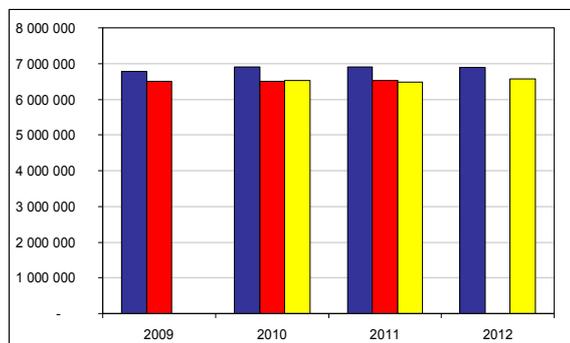
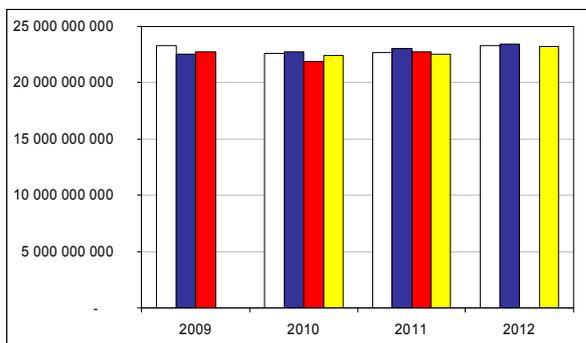
Champ : ménages ordinaires de France métropolitaine

Annexe 6 Comparaison de la distribution des niveaux de vie dans la population en 2011

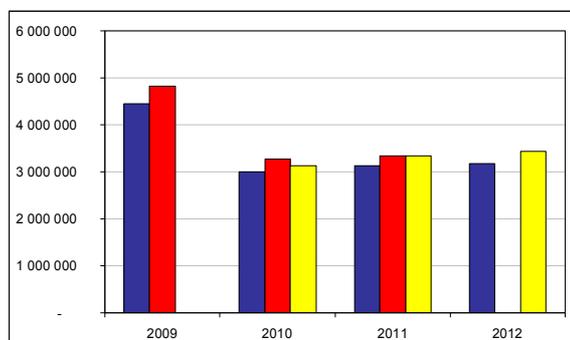
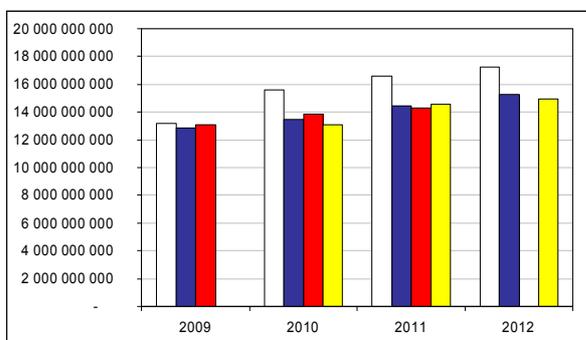


Champ : ménages ordinaires de France métropolitaine

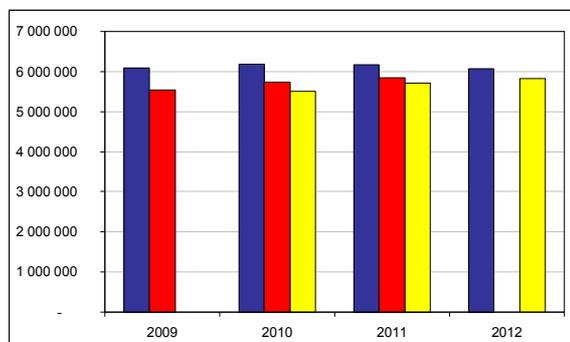
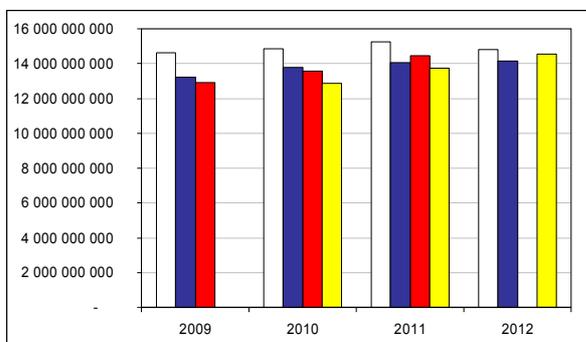
Annexe 7 Comparaison des principales composantes du revenu disponible
MASSES (€) **EFFECTIFS**



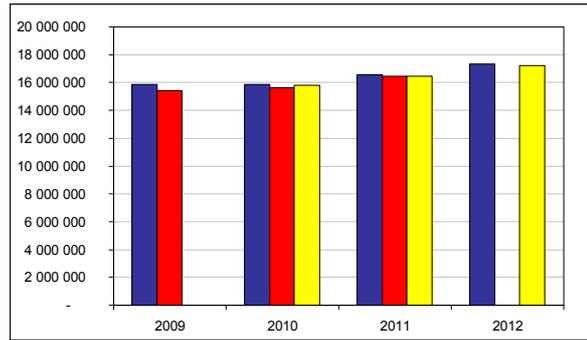
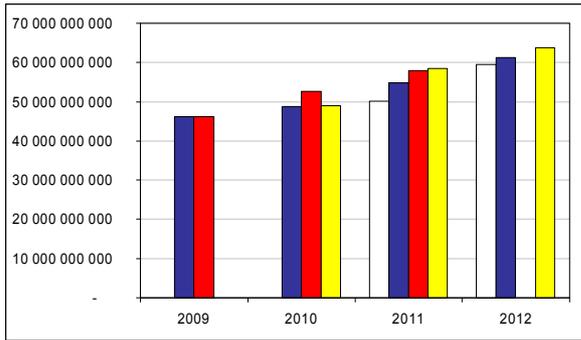
Prestations familiales



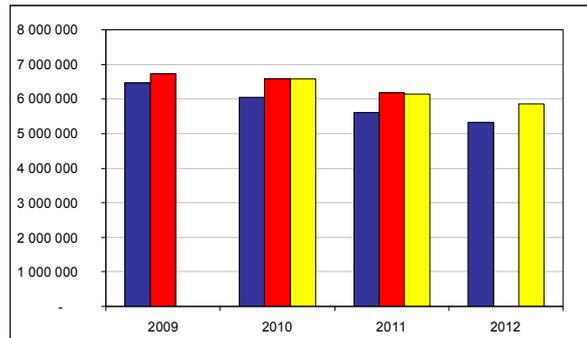
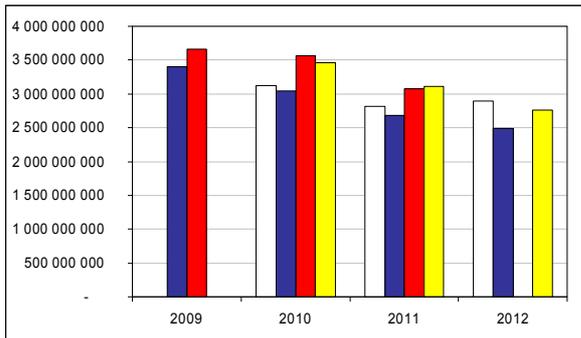
Minima sociaux (+ RSA activité)



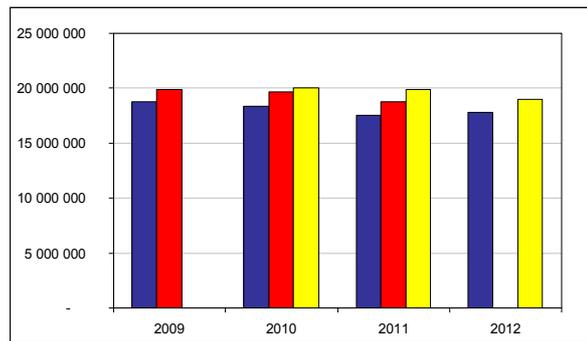
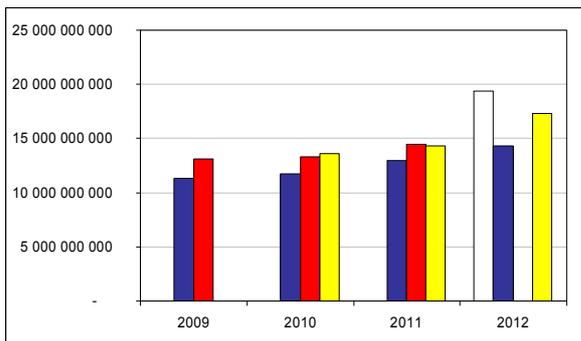
Allocations logement



Impôt sur le revenu (avant PPE)



Prime pour l'emploi

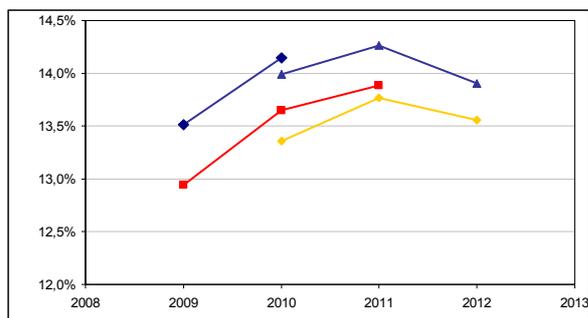


Prélèvements sur les revenus du patrimoine

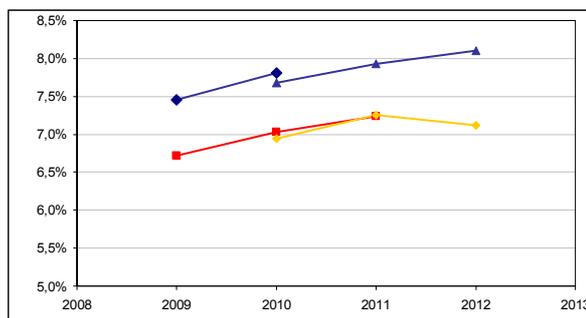
Champ : ménages ordinaires de France métropolitaine

Note : Les cibles sont issues de la Cnaf pour les prestations et des comptes des administrations publiques ou de la DGFiP pour les prélèvements. Elles sont corrigées par les soins des auteurs pour refléter plus fidèlement le champ de l'ERFS (ménages ordinaires de France métropolitaine).

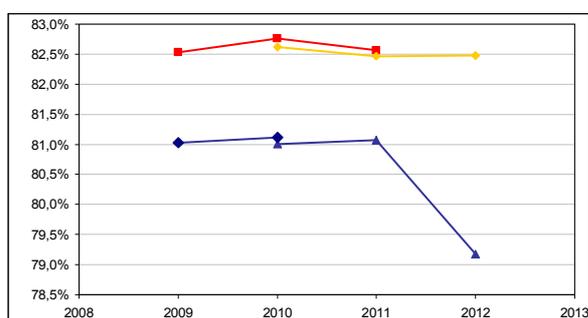
Annexe 8 Comparaison des principaux indicateurs de pauvreté et d'inégalités



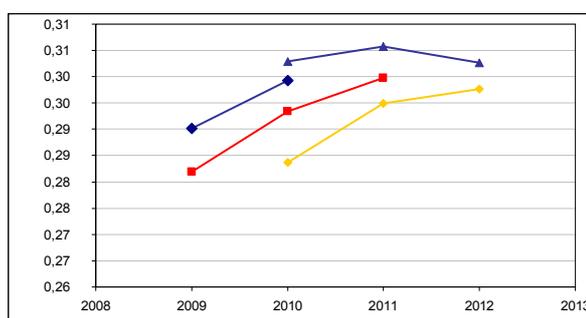
Taux de pauvreté à 60 %



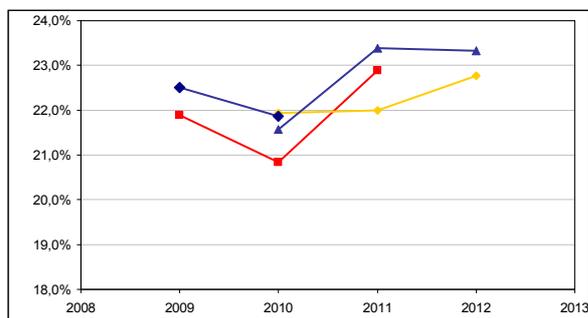
Taux de pauvreté à 50 %



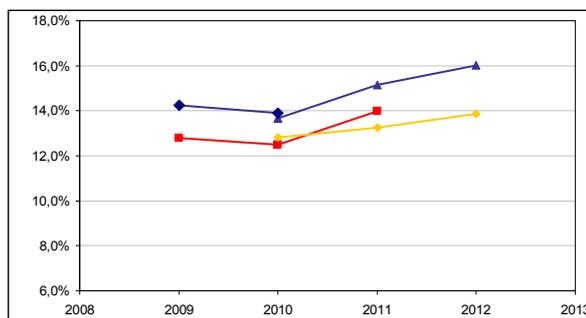
Intensité de la pauvreté à 60 %



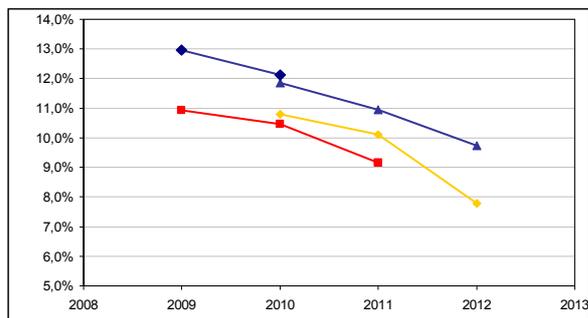
Indice de Gini



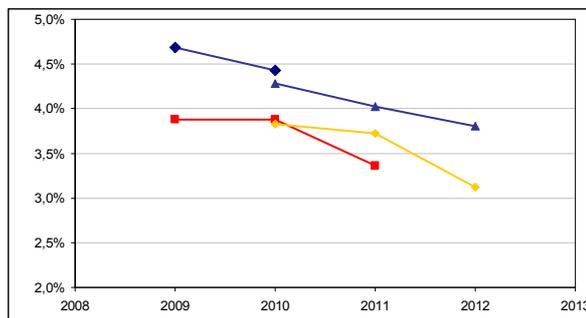
Taux de pauvreté des 18-24 ans à 60 %



Taux de pauvreté des 18-24 ans à 50 %



Taux de pauvreté des 75 ans et plus à 60 %

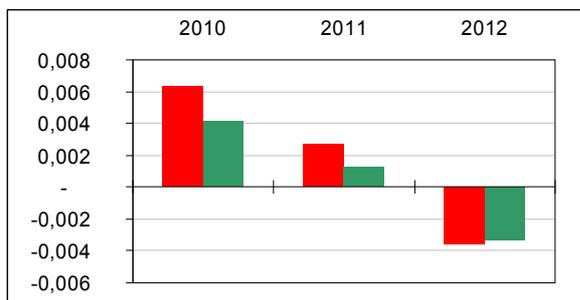


Taux de pauvreté des 75 ans et plus à 50 %

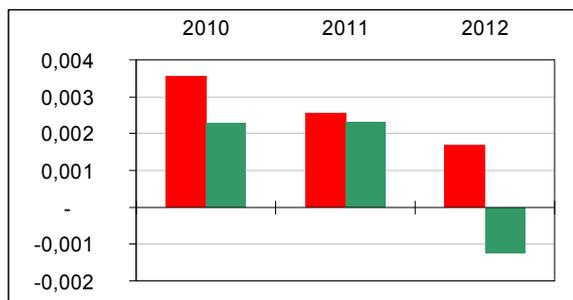
Champ : ménages ordinaires de France métropolitaine

Annexe 9 Comparaison des évolutions simulées et constatées

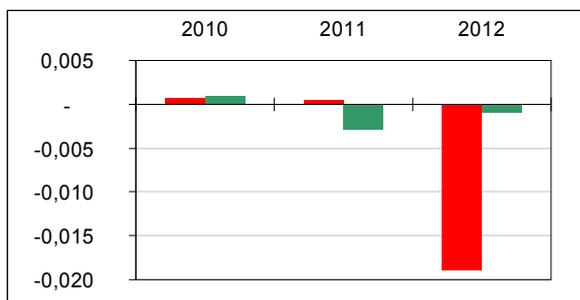
■ Evolution observée entre deux ERFS ■ Evolution simulée (Nowcasting)



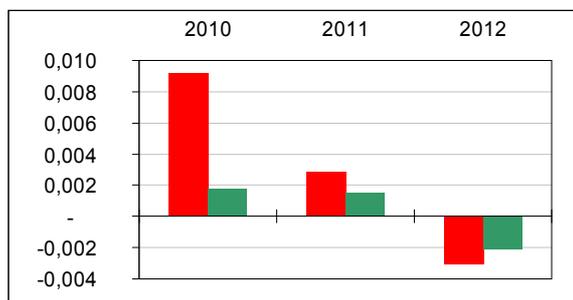
Taux de pauvreté à 60 %



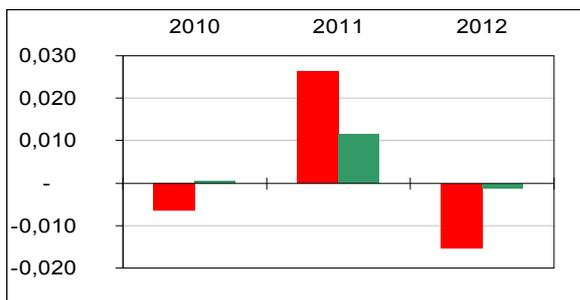
Taux de pauvreté à 50 %



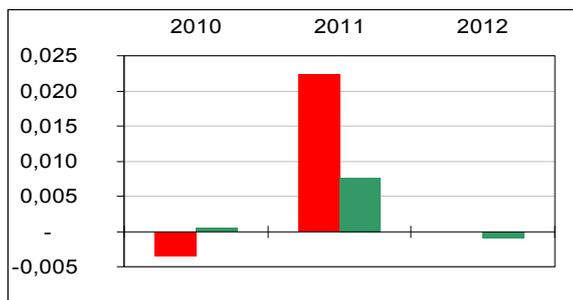
Intensité de la pauvreté à 60 %



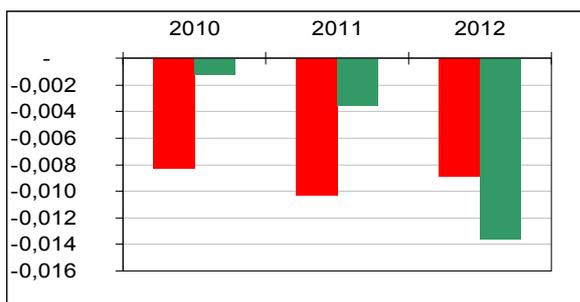
Indice de Gini



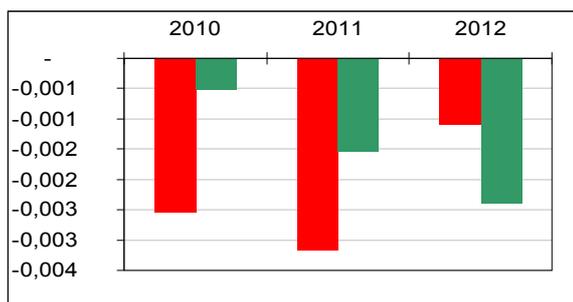
Taux de pauvreté des 18-24 ans à 60 %



Taux de pauvreté des 18-24 ans à 50 %



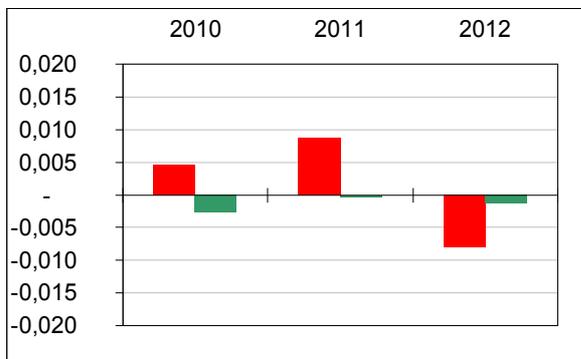
Taux de pauvreté des 75 ans et plus à 60 %



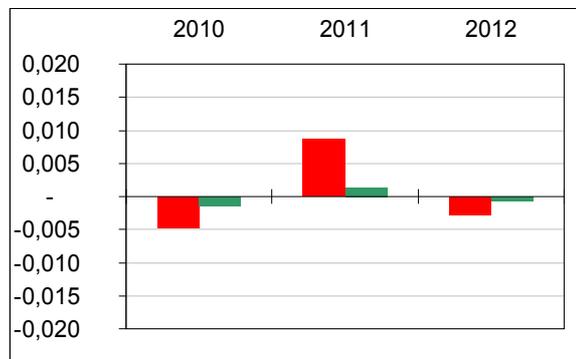
Taux de pauvreté des 75 ans et plus à 50 %

Annexe 10 Comparaison des évolutions simulées et constatées des taux de pauvreté à 60 % par type de ménage

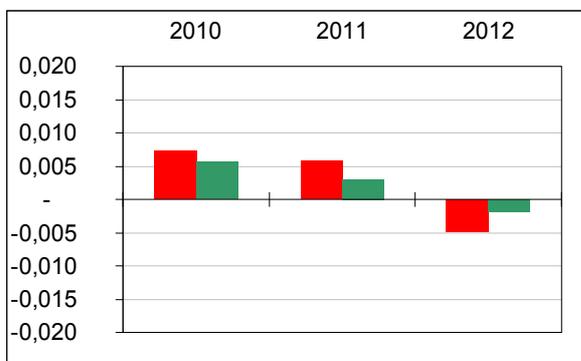
■ Evolution observée entre deux ERFS ■ Evolution simulée (Nowcasting)



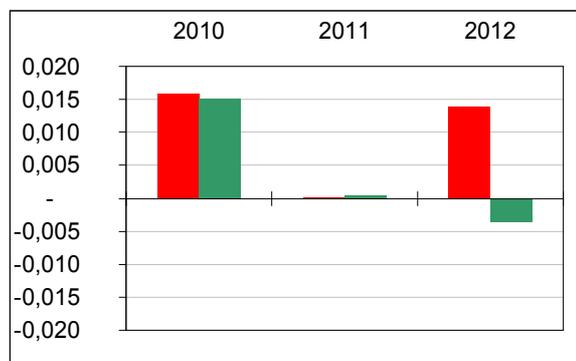
Personnes seules de moins de 65 ans



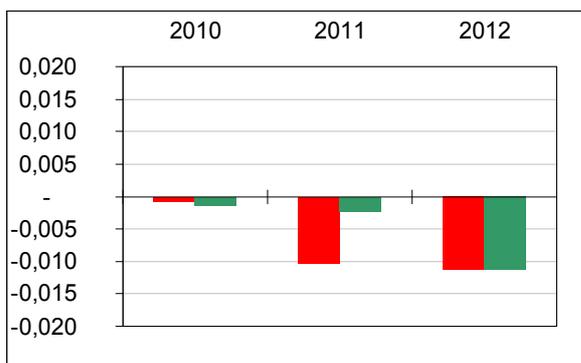
Couples sans enfants dont la personne de référence a moins de 65 ans



Couples avec enfants dont la personne de référence a moins de 65 ans



Familles monoparentales dont la personne de référence a moins de 65 ans

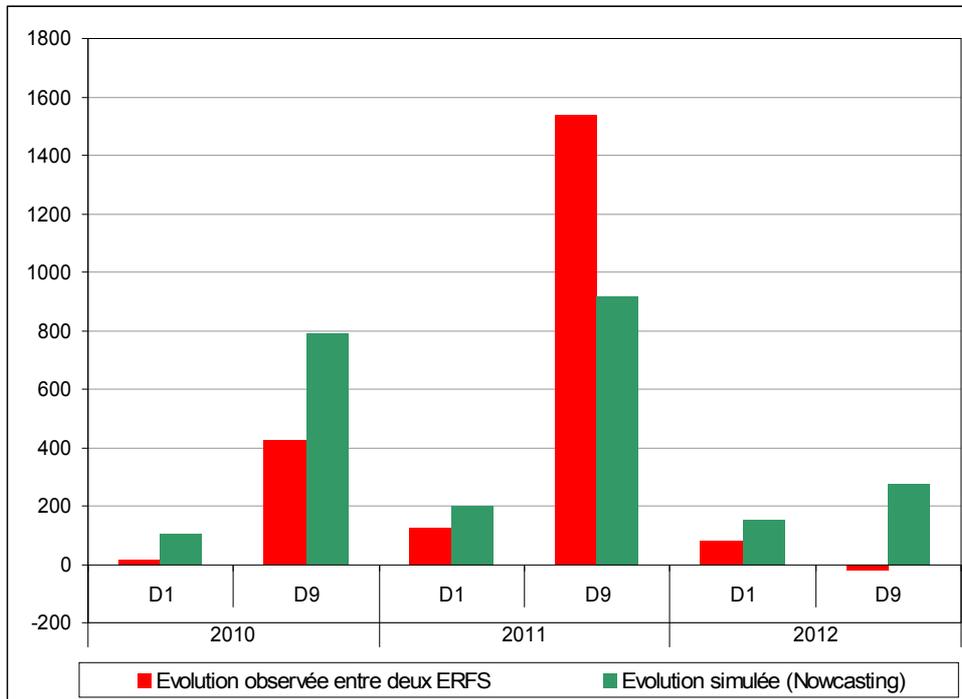


Ménages dont la personne de référence a 65 ans ou plus

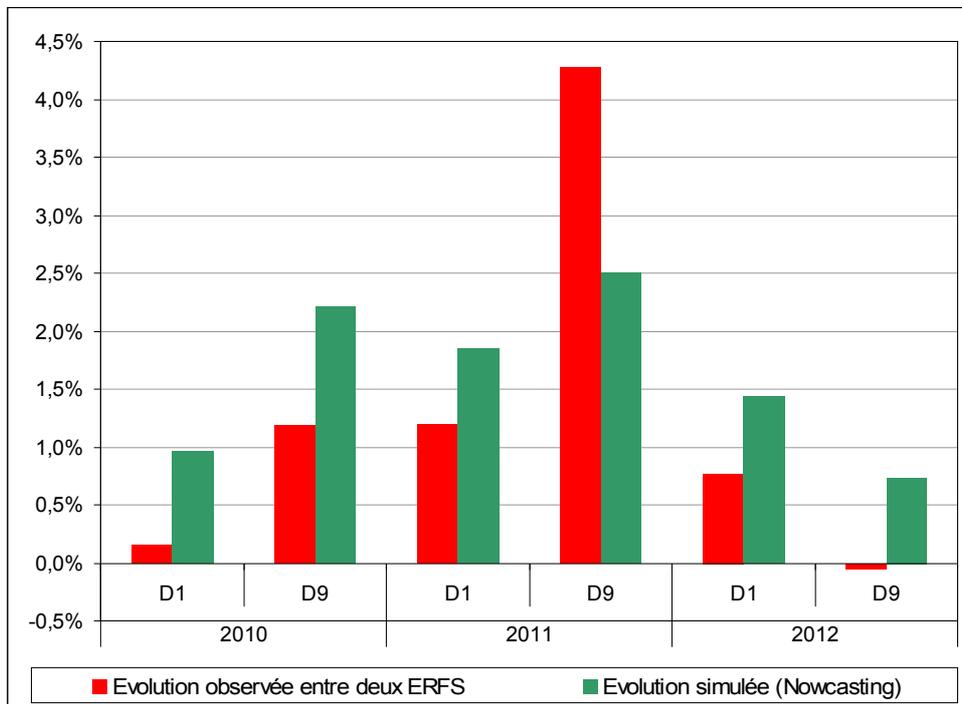
Annexe 11 Comparaison des évolutions simulées et constatées des premier et dernier décile de niveau de vie, par rapport à l'année précédente

En niveau (euros courants)

En



pourcentage d'évolution



Note de lecture : entre 2009 et 2010, nous aurions estimé une augmentation de 104 euros du premier décile de niveau de vie (figure du haut), soit +1,0 % (figure du bas), alors que d'après l'ERFS il a augmenté de 16 euros (figure du haut), soit +0,2 % (figure du bas).