

Rétropolation des poids dans les fichiers individuels de l'EEC en 2014-2020 suite à la refonte de l'enquête Emploi en 2021

Synthèse

- En 2021, l'enquête Emploi a connu une refonte majeure, qui entraîne une révision des principaux indicateurs issus de l'enquête. Des nouvelles séries longues ont été produites par l'Insee, à champ et mesure homogène depuis 1975.
- Dans le même temps, les marges démographiques sur lesquelles l'enquête est calée ont été elles-aussi fortement révisées, du fait de la prise en compte de l'impact de la nouvelle feuille de logement dans le Recensement.
- Sur la période 2014-2020, une pondération « rétrolée » est mise à disposition des utilisateurs des fichiers individuels de l'EEC. **Il est indispensable d'utiliser cette nouvelle pondération (EXTRI_RETROPOLE) pour estimer des indicateurs cohérents avec 2021.**
- Attention :
 - Il n'existe pas de pondération rétrolée avant 2014. Sur la période 1975-2013, les séries longues agrégées doivent être impérativement utilisées pour mesurer des évolutions à champ constant.
 - La pondération rétrolée 2014-2020 ne permet pas de corriger de l'ensemble des ruptures de série.

Contexte : une rupture de série en 2021 avec la refonte de l'enquête...

L'enquête Emploi a connu en 2021 [une refonte majeure](#). En effet, pour se conformer au nouveau règlement européen sur les statistiques sociales, entré en vigueur en début d'année, le questionnaire de l'enquête Emploi a été rénové. Dans le même temps, le protocole de l'enquête s'est modernisé : depuis le début de l'année, après une première interrogation maintenue en face-à-face, les ménages ont désormais la possibilité de répondre sur Internet en réinterrogation, en plus du téléphone. La méthode de pondération a également été revue.

Ces changements ont modifié le niveau des indicateurs issus de l'enquête. Pour pouvoir estimer au mieux les ruptures induites, une vaste opération méthodologique « Pilote » a été conduite tout au long de l'année 2020 et au premier trimestre 2021. Ceci a permis d'estimer de **nouvelles séries longues du marché du travail** « sans ruptures » sur les principaux indicateurs conjoncturels, afin de continuer d'éclairer la conjoncture du marché du travail sur longue période.

Au final, le *niveau* du taux d'emploi des 15-64 ans est relevé de 0,8 point en 2020, alors que le taux de chômage est inchangé ; la part de halo autour du chômage est rehaussée de 0,8 point. Les *évolutions* de l'ensemble des indicateurs corrigés de la rupture sont en revanche quasi inchangées sur le passé, notamment pendant la crise sanitaire et économique de 2020.

... et une forte révision des marges démographiques pour prendre en compte l'impact de la nouvelle feuille de logement du Recensement

Par ailleurs, une révision importante des marges démographiques sur lesquelles l'enquête Emploi est calée a été effectuée, pour prendre en compte l'impact de la [nouvelle feuille de logement et la refonte de l'analyse Ménages-Familles](#). Cette révision conduit à diminuer la population des personnes vivant en logement ordinaire de 0,6 % environ (*tableau 1*), l'impact étant beaucoup plus élevé chez les enfants et les jeunes (près de - 2 %). Les nouvelles séries longues du marché du travail sont mises en cohérence avec ces nouvelles populations.

Tableau 1 – Révision de la population par classe d'âge

	Moyenne 2014-2020		
	Ancienne pondération	Pondération rétrolée	Écart (en %)
Population totale	65 191 601	64 784 598598	-0,6
Moins de 15 ans	12 049 564	11 829 635	-1,8
15-24 ans	7 590 138	7 449 860	-1,8
25-29 ans	3 766 033	3 756 739	-0,2
30-49 ans	16 804 213	16 791 448	-0,1
50-64 ans	12 685 426	12 674 286	-0,1
65 ans ou plus	12 296 227	12 282 630	-0,1

La méthode de rétropolation

Les nouvelles séries longues ont été estimées pour les indicateurs principaux de l'enquête et par grands groupes de sous-populations (par tranche d'âge et sexe notamment), avec une méthode de rétropolation des données agrégées. Le principal objectif de la rétropolation des pondérations individuelles est de permettre à l'utilisateur de ces données de retrouver ces séries préalablement calculées, en agrégeant les données individuelles. Il est également utile de pouvoir calculer des données sur d'autres indicateurs de l'enquête ou sur d'autres sous-populations, qui n'ont pas fait l'objet de rétropolation agrégées. L'application d'un jeu de pondérations individuelles permet de conserver de la cohérence entre ces calculs sur les autres indicateurs de l'enquête et les indicateurs qui ont fait l'objet d'une rétropolation.

La méthode mise en place consiste à modifier les poids des années précédant la refonte de façon à ce qu'ils satisfassent des marges sur les niveaux rétropolés. Le principe de rétropolation est le même que celui opéré lors de la dernière refonte de l'enquête Emploi (2013), lorsque les pondérations individuelles pour 2003-2012 avaient été rendues cohérentes avec les nouvelles séries longues démarrant en 2013.

Pour le nouvel exercice, seuls les millésimes de 2014 à 2020 ont été rétropolés sur les nouvelles séries longues démarrant en 2021, du fait de la restriction de l'EEC à la France métropolitaine jusque 2013.

D'un point de vue pratique, un recalage a été effectué pour chaque trimestre entre le 1^{er} trimestre 2014 et le 4^e trimestre 2020, à partir des variables initiales de pondération (EXTRILOG au niveau logement et EXTRI au niveau individuel). Ces variables ont été recalées sur un ensemble d'indicateurs de façon à retrouver les mêmes résultats que pour les séries longues. La liste de ces indicateurs et le niveau de détail pris en compte pour les marges sont présentés dans le *tableau 2*.

Tableau 2 - Les indicateurs du marché du travail pris en compte dans la pondération rétropolée

Indicateur	Source	Calage par rang (variable RGA)	Niveau de détail
Statut BIT x âge x sexe	Nouvelles séries longues du marché du travail	Non	Âge quinquennal Emploi : Regroupé au-dessus de 75 et plus Chômage : Regroupé au-dessus de 65 et plus
Statut BIT par Dom x âge x sexe	Nouvelles séries longues du marché du travail	Non	Âge en 4 tranches (15-24, 25-49, 50-64, 65+)
Emploi par statut (<i>nouvelle variable STATUDET</i>)	Nouvelles séries longues du marché du travail	Non	Aucun
Emploi par groupe socioprofessionnel x sexe	Nouvelles séries longues du marché du travail	Non	PCS à 1 chiffre avec détail qualifié / peu qualifié pour employés et ouvriers
Temps partiel par âge x sexe	Nouvelles séries longues du marché du travail	Non	Âge en 3 tranches (15-24, 25-49, 50+)
Sous-emploi par âge x sexe x composante	Nouvelles séries longues du marché du travail	Non	Âge en 3 tranches (15-24, 25-49, 50+) Composante en 3 modalités
Halo par âge x sexe	Nouvelles séries longues du marché du travail	Non	Âge en 4 tranches (15-24, 25-49, 50-64, 65+)
Halo par composante	Nouvelles séries longues du marché du travail	Non	Composante en 3 modalités
Halo par Dom x âge x sexe	Nouvelles séries longues du marché du travail	Non	Âge en 4 tranches (15-24, 25-49, 50-64, 65+)
Chômage de longue durée par âge x sexe	Nouvelles séries longues du marché du travail	Non	Âge en 3 tranches (15-24, 25-49, 50+)
NEET par âge x sexe	Nouvelles séries longues du marché du travail	Non	Âge en 3 tranches (15-19, 20-24, 25-29)

Par ailleurs, comme indiqué précédemment, en même temps que la révision des indicateurs liés au marché du travail, une forte révision des marges démographiques est opérée, ce qui implique un calage concomitant avec de nouvelles marges démographiques (*tableau 3*).

Tableau 3 - Les indicateurs démographiques pris en compte dans la pondération rétropolée

Indicateur	Source	Calage par rang (variable RGA)	Niveau de détail
Population totale par âge x sexe	Démographie	Oui	Âge quinquennal Regroupé au-dessus de 75 ans et plus
Population totale par région	Démographie	Oui	Anciennes régions (26)

Les nouvelles variables mises à disposition

Chacune des tables trimestrielles de l'EEC pour 2014-2020 a au total été enrichie des variables de pondérations suivantes :

- EXTRI_RETROPOLE : pondération individuelle, ensemble des rangs, version rétropolée de la variable EXTRI
- EXTRILOG_RETROPOLE : pondération au niveau logement, ensemble des rangs, version rétropolée de la variable EXTRILOG
- EXTRID_RETROPOLE : pondération individuelle, rang 1, version rétropolée de la variable EXTRID
- EXTRIDF_RETROPOLE : pondération individuelle, rangs 1 et 6, version rétropolée de la variable EXTRIDF

À noter que les poids non rétropolés restent disponibles dans ces tables.

Par ailleurs deux variables ont été ajoutées :

- STATUTDET : statut d'emploi (variable mise en cohérence avec la nouvelle variable de statut disponible dans le FPR de l'EEC à partir de 2021)
 - o 10. Indépendant
 - o 21. CDI, fonctionnaire
 - o 22. CDD
 - o 23. Intérim
 - o 24. Alternance, stage
 - o 25. Sans contrat, ni stage
 - o 99. Non réponse
- NEET : ajout de la variable pour les millésimes où elle n'était pas disponible (2014 et 2015)

Précautions d'utilisation

Même si le nombre d'indicateurs pris en compte dans la réropolation de 2021 est important, il n'est pas possible de traiter l'ensemble des ruptures de séries induites par la refonte. D'une part, il n'est pas possible de distinguer chacune de ces ruptures, dont les causes peuvent être très variées (changement de questionnaire, impact du protocole ou du nouveau règlement, etc.). D'autre part, la prise en compte d'un nombre trop important d'indicateurs dans le calage entraînerait une trop forte proportion de poids élevés, ce qui nuirait à la fiabilité des résultats issus de la pondération réropolée.

L'utilisation de la pondération réropolée peut modifier les résultats pour des indicateurs et pour des sous-populations autres que ceux pris en compte explicitement comme marges. Par exemple, avec la refonte, on observe une hausse de la proportion des personnes se déclarant dans l'EEC inscrites à Pôle emploi. Cet indicateur n'est pas réropolé en tant que tel, mais son niveau augmente avec l'utilisation de la pondération réropolée, car il est très lié au halo autour du chômage (qui est lui réropolé en tant que tel). Mais il convient d'utiliser avec précaution les pondérations réropolées mises à disposition pour 2014-2020. En cas d'exploitation sur plusieurs années d'un indicateur non réropolé, il faut donc utiliser les poids réropolés et comparer les niveaux entre 2020 et 2021 pour vérifier qu'une rupture de série importante (non corrigée) n'a pas lieu. Autrement dit, l'utilisation des pondérations individuelles est une condition nécessaire pour obtenir des séries corrigées de rupture, mais pas suffisante : tout utilisateur doit également analyser si l'indicateur considéré ne comporte pas lui-même de rupture de série en 2021.

Enfin, il est important de noter que la réropolation n'a pas conduit à modifier les caractéristiques au niveau individuel, mais seulement les pondérations : les calculs effectués par exemple sur les transitions individuelles entre statuts d'activité, dont les concepts ont été redéfinis avec la refonte, ne sont plus pertinentes si l'on compare des périodes avant et après 2021 (et ce quelle que soit la pondération).

Résultats de la rétopolation

Le *tableau 4* recense quelques indicateurs mesurés en ancienne et en nouvelle pondération, et dont les séries longues ont été rétopolées.

La refonte de l'enquête Emploi de 2021 conduit à réviser à la hausse le taux d'emploi, en particulier pour les femmes et les seniors, ainsi que le halo autour du chômage. Elle contribue également à réviser nettement à la baisse le taux de chômage de longue durée. En revanche le taux de chômage n'est pas affecté, hormis celui des jeunes que la refonte conduit à réviser à la hausse.

Parmi les personnes en emploi, la part des professions intermédiaires et des artisans, commerçants et chefs d'entreprise est révisée à la baisse. La part des personnes à temps partiel est quant à elle révisée à la hausse. Pour disposer d'éléments d'explication sur l'impact de la refonte sur certains indicateurs, l'utilisateur peut se reporter à l'[Insee Analyses n°65](#) (2021).

Tableau 4 : Résultats de la rétopolation sur les principaux indicateurs du marché du travail

	Moyenne 2014-2020 (en %)		
	Avant rétopolation	Après rétopolation	Écart (en point)
Taux d'emploi			
15-64 ans	64,7	65,5	0,8
15-64 ans, hommes	68,0	68,3	0,2
15-64 ans, femmes	61,4	62,8	1,4
15-24 ans	28,6	29,1	0,4
25-49 ans	80,1	80,7	0,6
50-64 ans	61,1	62,2	1,1
65-74 ans	4,9	5,0	0,0
Taux de chômage	9,4	9,4	0,0
Hommes	9,5	9,5	0,0
Femmes	9,2	9,2	0,0
15-24 ans	22,3	23,6	1,4
25-49 ans	8,8	8,6	-0,2
50 ans ou plus	6,5	6,5	0,0
Taux de chômage de longue durée	4,0	2,7	-1,3
Halo au tour du chômage			
Part parmi les 15-64 ans	3,9	4,5	0,7
Part dans l'emploi			
Agriculteurs	1,6	2,0	0,4
Artisans, commerçants, chefs d'entreprise	6,5	5,9	-0,6
Cadres	18,3	19,0	0,7
Professions intermédiaires	25,7	24,8	-0,9
Employés	27,3	28,0	0,7
Ouvriers	20,1	19,9	-0,2
Indépendants	11,8	11,9	0,1
CDI, fonctionnaires	74,4	74,6	0,2
CDD	8,5	8,2	-0,3
Intérim	2,3	2,0	-0,3
Alternance, stage	2,4	2,4	0,0
Sans contrat, indéterminé	0,5	0,9	0,4
Temps partiel	18,5	19,2	0,6
Sous-emploi	6,9	7,2	0,3

La méthode de rétopolation par les poids suppose l'utilisation de nombreuses marges de calage supplémentaires aux marges sociodémographiques usuelles. Malgré les nombreuses contraintes de calage supplémentaires imposées, la rétopolation des poids ne conduit pas à une dispersion des poids trop importante (tableau 5) :

- 99 % des poids rétopolés sont inférieurs à 1 883, contre 1 801 pour les poids avant rétopolation
- Aucun poids rétopolé n'est supérieur à 10 000
- Dans 80 % des cas, le ratio entre poids après/avant rétopolation est compris entre 0,8 et 1,2

Tableau 5 – comparaison des distributions des deux jeux de poids

	Poids avant rétopolation	Poids après rétopolation	Ratio après/avant
Minimum	1	1	1,00
10 ^e percentile	171	161	0,80
1 ^{er} quartile	342	326	0,91
Moyenne	518	515	0,99
3 ^e quartile	626	626	1,05
90 ^e percentile	834	858	1,19
99 ^e percentile	1 801	1 883	1,70
Maximum	12 708	9 493	13,60