

Annexe 2 : une méthode possible pour un correctif macro de la pyramide des âges

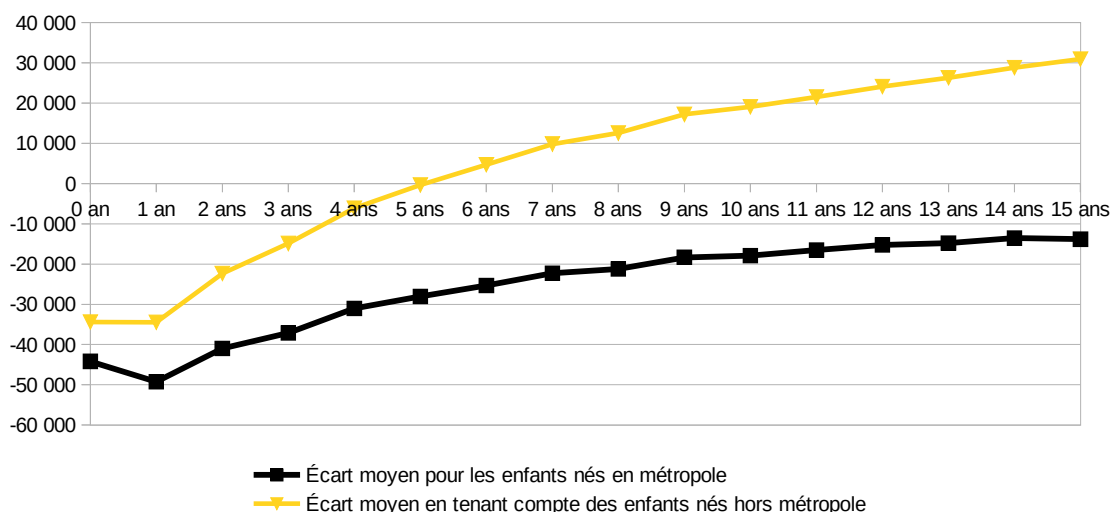
Les chiffres présentés dans le tableau 3-B.1 de la partie 3-B « Omissions possibles de jeunes enfants au recensement : que nous apprennent les comparaisons avec l'état civil ? » concernent les enfants nés en métropole, pour disposer des données nécessaires sur le même champ géographique pour toutes les années pour lesquelles la comparaison entre l'EAR et l'état civil a été réalisée. Dans le même temps, des enfants nés hors métropole arrivent chaque année en France. S'ils ne figurent pas dans les fichiers de naissances de l'état civil³², ils sont comptabilisés au recensement. Celui-ci fournit le nombre d'enfants nés hors métropole et vivant en métropole à un âge donné. Pour chaque âge entre 0 et 15 ans, on calcule dans chaque EAR le nombre d'enfants recensés nés hors France métropolitaine. Puis, pour chaque âge, on en calcule la moyenne sur l'ensemble des EAR. Ainsi, en moyenne sur les EAR de 2004 à 2018, 10 000 enfants nés hors métropole sont recensés en métropole à 0 an. Leur nombre augmente avec l'âge : ils sont 45 000 en moyenne à 15 ans.

On peut alors considérer que ces enfants nés hors métropole viennent « compenser » au recensement, du point de vue du « nombre de personnes d'un âge donné présent en France », une partie des omissions de jeunes enfants nés en métropole, comme en atteste le graphique A2.1.

En tenant compte des enfants nés hors métropole, l'écart entre le recensement et l'état civil devient nul à partir de 5 ans. Au-delà de cet âge, le recensement comptabilise davantage d'enfants que l'état civil. Ainsi, l'augmentation dans les EAR du nombre d'enfants nés à l'étranger ou dans les DOM et les COM et vivant en métropole permet de compenser vers 5 ans le déficit observé sur les enfants nés en métropole. Il n'en reste pas moins que, sur les seuls nés en métropole, lorsque l'on compare les chiffres des EAR à ceux de l'état civil, il reste un déficit plus important et plus persistant, mais dont une partie au moins provient des départs d'enfants nés en France et vivant à l'étranger.

32 C'est en particulier le cas des enfants nés à l'étranger.

Graphique A2.1 : Avant 10 ans, les écarts entre le recensement et l'état civil se réduisent avec l'âge et avec la prise en compte des enfants nés hors métropole



Champ :

pour la courbe sur les enfants nés en métropole : enfants recensés et nés en France métropolitaine – Naissances enregistrées en métropole, diminuées des décès enregistrés en métropole ;

pour la courbe tenant compte des enfants nés hors métropole : enfants recensés en France métropolitaine quel que soit leur lieu de naissance – Naissances enregistrées en métropole, diminuées des décès enregistrés en métropole.

Source : Insee, *Enquêtes annuelles de recensement de 2004 à 2018 et statistiques d'état civil*.

Dans le cadre de ses travaux sur la comparaison entre les effectifs de naissances de l'état civil et les effectifs des jeunes enfants au recensement, l'Insee a élaboré les grands principes que pourrait suivre un correctif macro de la pyramide des âges si l'on souhaitait corriger le sous-dénombrement des jeunes enfants au recensement. Il s'agirait, à un niveau macro, de modifier le nombre d'individus âgés de 0 à 5 ans, en supposant que les effectifs à 6 ans ne sont plus sous-estimés au recensement (cf. graphique A2.1). Les effectifs à 0 an pourraient être estimés, à partir des données d'état civil, par les naissances diminuées des décès. Les effectifs de 1 à 5 ans pourraient être obtenus par interpolation linéaire entre ces effectifs estimés à 0 an et les effectifs, estimés ou observés dans les EAR, à 6 ans : la méthode repose ainsi sur l'existence d'une fonction affine reliant les effectifs d'une même génération entre 0 et 6 ans et sur le suivi d'une génération sur plusieurs enquêtes de recensement successives, ce qui a du sens d'un point de vue démographique.

Après avoir estimé ces nouveaux effectifs pour chaque EAR, les pyramides des âges pourraient être construites selon la méthode habituelle : pyramide des âges calculée par génération, calée sur le nombre total d'habitants du recensement (RP cumulé), comme cela est fait habituellement.

Il s'agirait donc bien d'un correctif de niveau macro qui ne modifierait en rien les données individuelles des EAR, y compris notamment le chiffre connu et définitif de la population française donné par le dernier recensement.

Estimation des effectifs à 0 an

L'estimation des effectifs à 0 an repose sur les données d'état civil. Les effectifs à 0 an dans l'EAR de l'année N sont estimés par les naissances enregistrées de l'année N-1, auxquelles on soustrait les décès enregistrés à 0 an en N-1, et ce pour chacun des deux sexes. Pour l'EAR 2009 par exemple, les effectifs (notés *Eff* dans la formule ci-dessous) à 0 an au 1^{er} janvier seront estimés comme suit :

$$\text{Eff 0 an EAR 2009}^* = \text{0 an génération 2008 au 1}^{\text{er}} \text{ janv 2009} = \text{Naissances enregistrées en 2008} - \text{Décès enregistrés à 0 an en 2008}$$

* : données estimées

Eff : Effectifs

Les naissances et décès enregistrés (naissances en France) ont été préférés aux naissances et décès domiciliés (personnes vivant en France) puisque le redressement porte sur les effectifs globaux des EAR et non uniquement sur les enfants nés en France. Par ailleurs, le fait de ne pas exclure les enfants nés en France de mère domiciliée à l'étranger permet de « compenser » une petite part des enfants nés à l'étranger mais recensés sur le territoire français. Ce choix est en adéquation avec l'objectif fixé de raisonner en déficit « apparent » : on cherche à évaluer l'écart entre les effectifs à 0 an dans les EAR et les naissances (diminuées des décès), indépendamment du lieu de naissance.

Pour estimer les effectifs à 0 an, on ne tient pas compte à ce stade du solde migratoire. Le solde migratoire apparent à 0 an, calculé par différence entre deux pyramides des âges au 1^{er} janvier de deux années successives diminuées des décès à cet âge, est très négatif, du fait de la sous-estimation des bébés au recensement. Le seul moyen de réconcilier les données des naissances de l'état civil et le bas de la pyramide des âges issue du recensement (ou EAR) est de faire partir de nombreux bébés nés en France, d'où un solde très négatif à cet âge. On ne tient donc pas compte de ce solde pour estimer le point de départ des 0 an, ce qui revient, à ce stade du calcul, à faire l'hypothèse que le solde migratoire est nul. On fait cette hypothèse de travail parce qu'aucune donnée ne permet de l'estimer. Pour autant, le solde final ne sera pas nul lorsqu'il sera calculé par différence des pyramides corrigées (cf. graphique A 2.4 en annexe).

Interpolation linéaire des 1-5 ans dans les EAR

Le principe de l'interpolation linéaire est de calculer une droite affine entre deux points déterminés, ici 0 et 6 ans. Les effectifs à 0 an sont estimés comme mentionné ci-dessus. Les effectifs à 6 ans sont ceux de l'EAR pour la génération considérée.

Pour illustrer la méthode de manière plus concrète, on se place ici dans le cas où l'on souhaite établir la **pyramide des âges au 1^{er} janvier 2013**. Cette pyramide est construite à partir de cinq EAR, de 2011 à 2015. Les effectifs des 1-5 ans sont interpolés, par sexe, dans chacune de ces cinq EAR.

Pour déterminer, par exemple, les effectifs âgés de 2 ans corrigés dans l'EAR 2011, le calcul par interpolation linéaire repose sur les effectifs de cette génération à 0 an (que l'on a estimés) et à 6 ans (que l'on constate dans l'EAR 2015). Le calcul s'effectue de cette manière :

$$\text{Eff 2 ans EAR 2011}^* = \text{0 an génération 2008 au 1}^{\text{er}} \text{ janv 2009}^* + 2 \left(\text{0 an génération 2008 au 1}^{\text{er}} \text{ janv 2009}^* - \text{Eff 6 ans EAR 2015} \right) / 6$$

* : données estimées

Eff : Effectifs

De la même manière, les effectifs âgés de 5 ans dans l'EAR 2013 s'obtiennent *via* la formule :

$$\text{Eff 5 ans EAR 2013}^* = \text{0 an génération 2007}^* \text{ au 1}^{\text{er}} \text{ janv 2008} + 5 \left(\text{0 an génération 2007}^* \text{ au 1}^{\text{er}} \text{ janv 2008} - \text{Eff 6 ans EAR 2014} \right) / 6$$

* : données estimées

Eff : Effectifs

On voit ici que les effectifs des enfants de 6 ans non encore connus dans les EAR (car la génération n'a pas encore 6 ans dans l'EAR la plus récente, elle atteindra cet âge dans des EAR à venir dans le futur) posent un problème qui sera traité au paragraphe suivant. Par ailleurs, on notera qu'il est théoriquement possible que les effectifs des EAR pour un âge donné soient supérieurs aux effectifs obtenus par interpolation. Dans ce cas, on choisit de retenir les effectifs interpolés.

Une fois le redressement des EAR ainsi effectué, on procède au calcul de la pyramide des âges comme fait habituellement.

Estimation des effectifs à 6 ans pour les générations qui n'ont pas encore été recensées à cet âge

L'interpolation linéaire telle que présentée ci-dessus ne peut être appliquée directement pour les plus jeunes générations, car leurs effectifs à 6 ans ne sont pas encore connus. Par exemple, les enfants de 1 an dans la dernière EAR n'ont nécessairement pas encore été recensés à 6 ans et ne le seront que dans cinq ans. Il va donc falloir les estimer.

Pour estimer la population à 6 ans qui n'est pas encore connue, on va utiliser le rapport moyen entre la population recensée à 6 ans et celle estimée à 0 an (c'est-à-dire les naissances diminuée des décès) observé sur des générations pour lesquelles on dispose de toute l'information nécessaire.

On se place à nouveau dans le cadre du calcul de la pyramide des âges au 01/01/2013. Le tableau A2.1 montre comment les effectifs à 6 ans des générations 2009 à 2012 sont estimés à partir du rapport moyen entre les enfants de 6 ans et les 0 an observé sur les trois générations précédentes, les générations 2006 à 2008. La moyenne du rapport entre population recensée à 6 ans et population estimée à 0 an sur ces trois générations est de 1,0062 pour les garçons et de 1,0003 pour les filles. Appliquée aux effectifs à 0 an, elle permet d'estimer les effectifs à 6 ans des générations 2009 à 2012.

Tableau A2.1. Estimation des effectifs d'hommes et de femmes âgés de 6 ans des générations 2009 à 2012 à partir de l'évolution moyenne entre 0 an et 6 ans des générations 2006 à 2008

Année de naissance	Hommes			Femmes		
	0 estimé = N-D	6 ans (recensés)	6 / (N-D)	0 estimé = N-D	6 ans (recensés)	6 / (N-D)
2006	406 394	407 040	1,0016	387 979	388 161	1,0005
2007	400 859	404 264	1,0085	382 689	387 461	1,0125
2008	405 399	408 906	1,0086	388 143	383 518	0,9881
Moyenne de 6 / (N-D)			1,0062			1,0003
Année de naissance	0 estimé = N-D	6 ans estimés = (N-D) * 1,0062		0 estimé = N-D	6 ans estimés = (N-D) * 1,0003	
2009	404 453	406 979		386 403	386 535	
2010	408 781	411 334		390 986	391 120	
2011	403 971	406 494		386 772	386 904	
2012	403 463	405 983		384 506	384 637	

Champ : France métropolitaine

Quand il s'agira de calculer la pyramide des âges au 01/01/2014, la moyenne du rapport 0-6 ans sera calculée sur les générations 2007 à 2009 et permettra d'estimer les effectifs à 6 ans des générations 2010 à 2013.

Il faut noter qu'à l'arrivée d'une nouvelle EAR, on recalcule les effectifs corrigés des EAR précédentes, car de nouveaux effectifs à 6 ans sont désormais connus.

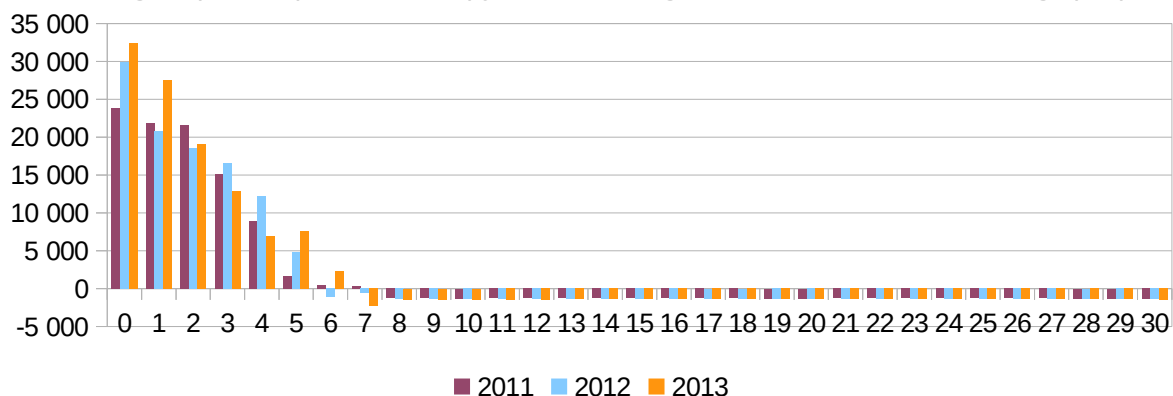
Résultats

On présente ici l'impact de la méthode de redressement sur 3 pyramides des âges pour la France métropolitaine, aux 1^{ers} janvier 2011, 2012 et 2013.

Par rapport aux pyramides actuellement diffusées, les modifications de population concernent principalement, et logiquement, les individus âgés de 0 à 5 ans (cf. graphique A2.2). L'écart le plus important s'observe toujours à l'âge 0 : par exemple, le correctif ajoute 32 400 enfants de 0 an à la pyramide au 01/01/2013. Cela représente 4,3 % de l'effectif initial des 0 an de cette année. De 0 à 5 ans, les écarts sont globalement décroissants avec l'âge.

En dehors de ces âges, la population des « pyramides interpolées » est, en moyenne sur les trois pyramides, légèrement inférieure d'environ 1 400 individus à chaque âge entre 8 ans et 65 ans (de moins en moins au-delà). En effet, la population municipale cible étant inchangée (c'est le total de la population du recensement), l'ajout de jeunes enfants dans les « pyramides interpolées » se traduit par une réduction du nombre d'individus aux âges plus élevés. Les écarts entre la pyramide de départ et la pyramide corrigée pour les âges intermédiaires, 6 et 7 ans, varient d'une année à l'autre, mais restent limités.

Graphique A2.2 : Écarts de population entre les pyramides des âges aux premiers janvier 2011, 2012 et 2013 corrigées par interpolation et les pyramides des âges définitives des Bilans démographiques.



Le tableau A2.2 reprend ces éléments en proposant une comparaison chiffrée des écarts entre les pyramides interpolées et les pyramides des âges non corrigées (i.e. celles diffusées).

Tableau A2.2 : Comparaison selon différents critères des pyramides interpolées et des pyramides actuelles

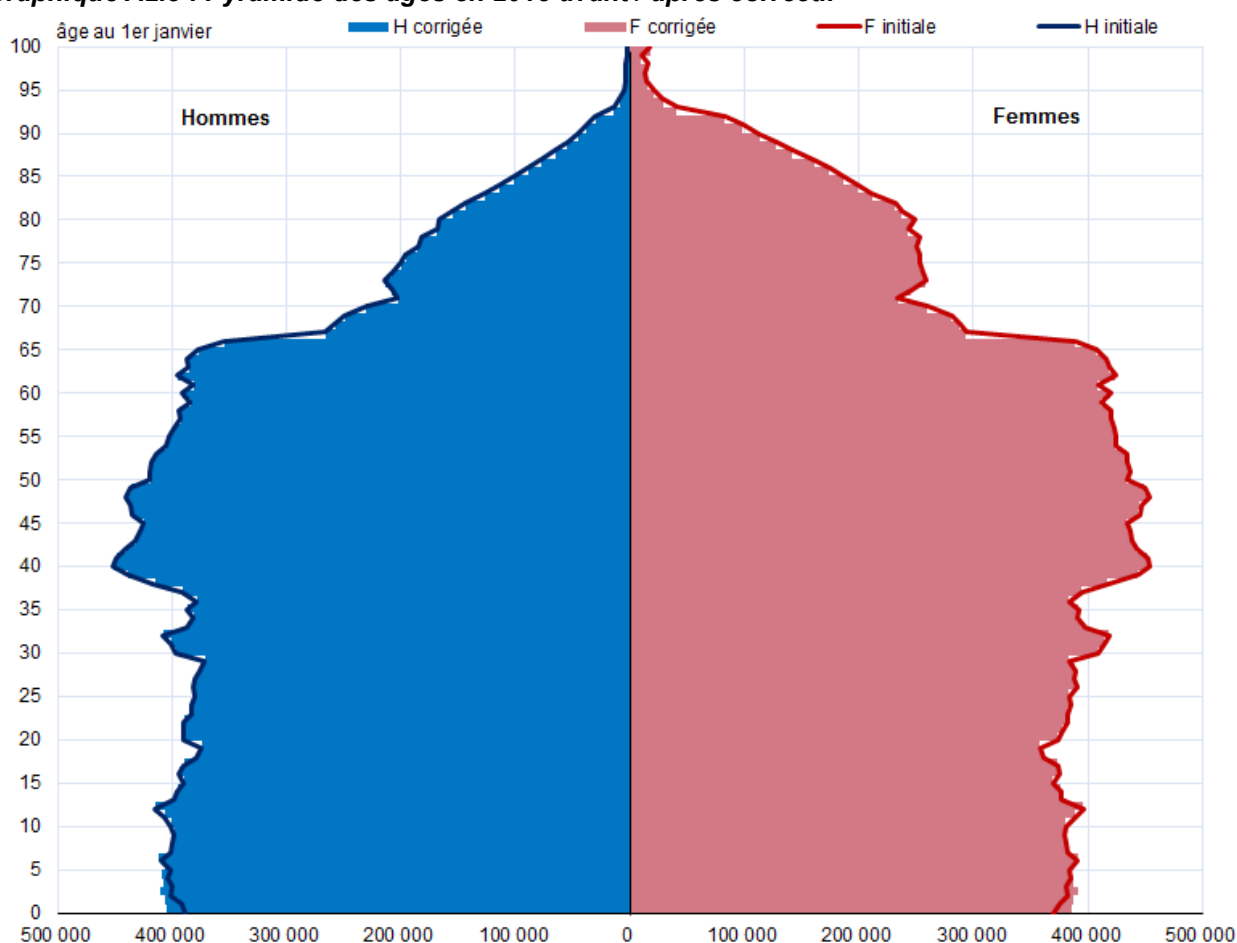
	Pyramide au		
	01/01/2011	01/01/2012	01/01/2013
0 an			
Écart absolu	+ 23 800	+ 29 800	+ 32 400
Écart relatif*	3,1 %	3,9 %	4,3 %
0-5 ans			
Écart absolu total	+ 92 900	+ 102 700	+ 106 400
Écart relatif**	2,0 %	2,2 %	2,3 %
8-65 ans			
Écarts moyens à chaque âge	- 1 300	- 1 400	- 1 500

* Écart relatif calculé par rapport aux effectifs à 0 an de la pyramide actuelle.

** Écart relatif calculé par rapport aux effectifs des 0-5 ans de la pyramide actuelle.

Comme attendu, l'interpolation linéaire corrige significativement le bas de la pyramide des âges. Il en découle qu'elle résout aussi au niveau national le problème majeur du solde migratoire fortement négatif à 0 an³³ et diminue le solde migratoire aux jeunes âges. Le solde migratoire à 0 an est estimé entre - 35 000 et - 40 000 (les sorties excèdent les entrées), surestimé avec les pyramides des âges, est également nettement réduit avec les pyramides obtenues par interpolation : il varie autour de - 6 000.

Graphique A2.3 : Pyramide des âges en 2013 avant / après correctif



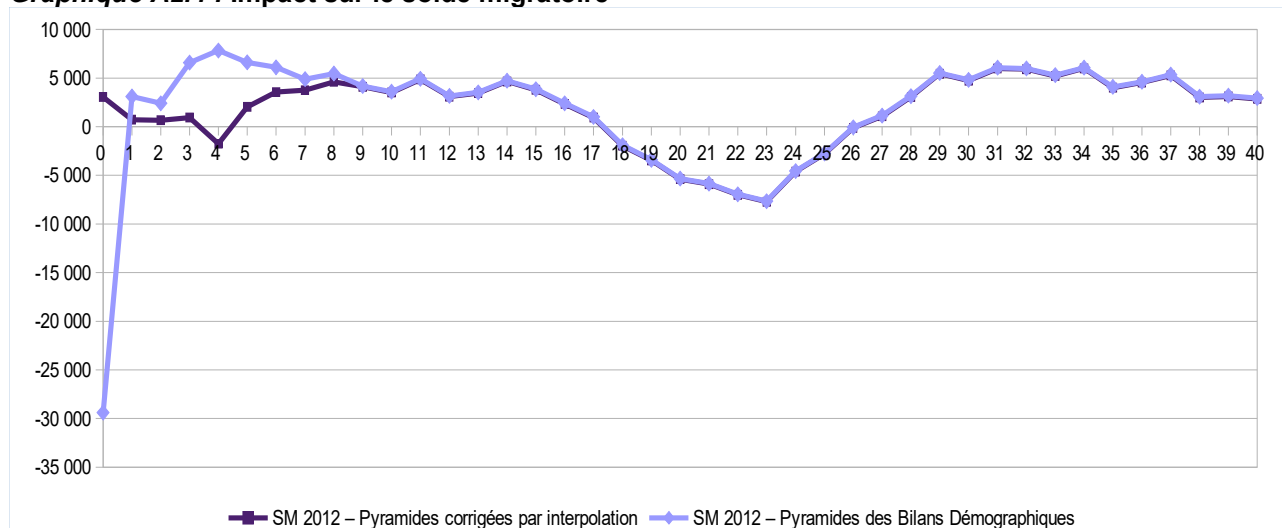
Champ : France métropolitaine.

Source : Insee.

33 La sous-estimation des enfants de 0 an au recensement et le mode de calcul du solde migratoire par différence entre deux pyramides des âges successives impliquent nécessairement un solde migratoire à 0 an fortement négatif quand il est déduit des pyramides des âges actuelles.

Le solde migratoire est déterminé à partir de la différence de deux pyramides des âges successives³⁴, après prise en compte de la natalité et de la mortalité. Ci-après figure le solde migratoire avant et après correction des pyramides des âges.

Graphique A2.4 : Impact sur le solde migratoire



Champ : France métropolitaine.

Source : Insee.

Sous les hypothèses retenues, l'interpolation linéaire est une méthode relativement simple et rapide à mettre en œuvre au niveau macro au niveau national. Elle peut rendre service aux utilisateurs des pyramides des âges qui seraient mis en difficulté par l'impact, sur ces pyramides, du déficit des jeunes enfants tel que mesuré au recensement. Elle nécessite des investigations plus importantes pour une déclinaison par région ou département. Elle n'est pas envisageable à un niveau géographique infra-départemental, du fait de la méthode d'échantillonnage du recensement de la population.

34 Le solde migratoire apparent à 0 an est quant à lui obtenu par la différence entre les effectifs à 0 an et les naissances de l'année précédente corrigée des décès à 0 an de cette même année.