
CONSOLIDER LES RÉPONSES DES UNITÉS LÉGALES POUR UNE STATISTIQUE D'ENTREPRISES PLUS COHÉRENTE

Guillaume CHANTELOUP

Insee, Direction des Statistiques d'Entreprises

guillaume.chanteloup@insee.fr

Mots-clés : Statistique d'entreprises, unité statistique, gestion d'enquête, source administrative

Résumé

Dans les statistiques d'entreprises, le profilage a déjà commencé à être pris en compte, mais uniquement pour les plus grands groupes, qui sont profilés manuellement. Pour les groupes restants, un algorithme est mis en place. Le profilage dit automatique consiste à considérer le groupe comme une seule entreprise, et estimer les flux intra sur le chiffre d'affaires à partir d'hypothèses. Les réponses aux enquêtes et autres données fiscales utilisées dans l'élaboration des statistiques annuelles d'entreprises sont obtenues par une méthode bottom-up : les valeurs des variables de l'entreprise sont obtenues en sommant les valeurs des variables des Unités Légales (UL) de son périmètre. Pour les variables non additives liées au chiffre d'affaires, on retranche un montant de flux intra-groupe estimé à partir des caractéristiques (Ape, chiffre d'affaires et achats) des UL.

La consolidation automatique conduit globalement à retirer 7 % du chiffre d'affaires de l'ensemble des unités légales qui composent les entreprises.

Une confrontation avec des données consolidées issues des groupes eux-mêmes montre que dans la majorité des cas, l'algorithme de consolidation consolide le montant exact ou consolide moins, ce qui est la cible retenue a priori, quel que soit le secteur de l'entreprise. Quoi qu'il en soit, une consolidation même prudente reste toujours plus proche de la réalité que la somme simple des UL qui composent l'entreprise et une sous-consolidation permet d'éviter de supprimer de l'activité économique à tort.

L'entreprise devenant l'unité statistique pour la statistique structurelle, c'est elle qui est l'unité à contrôler lors du processus ESANE et non plus les unités légales du contour même si ces dernières demeurent les unités de collecte. Ainsi, les macro-contrôles ne feront ressortir que des dossiers d'entreprises profilées. Les gestionnaires qui validaient uniquement jusqu'à présent les comptes des UL vont devoir contrôler dans un premier temps les UL de contour si elles ressortent à traiter, et le flux de l'entreprise dans un second temps. Cette nouvelle méthode de contrôle des données est présentée dans la dernière partie de cet article.

Abstract en anglais

The automatic profiling method provides a consolidated account, with a list of hypothesis. The variables used in the Esane application at the enterprise level are obtained from a bottom-up method : the value of the variables are summed from the variables of the Legal Unit in the perimeter. For non-additive variables such as turnover, an estimated amount of intra-flow is withdrawn.

The consolidation effect results in a reduction of about 7% in total turnover of the LEUs that make up target population. The effect of the new distribution of turnover throughout the whole economy

after consolidation remains modest. The confrontation with the data consolidated by the groups from a survey, proves that the algorithm does not consolidate too much, and is often careful.

Now, this new method forces the clerks to control the variables of the enterprises. If an enterprise has to be controlled, the clerks have to check the variables of the legal units, and if it is a non-additive variable, the amount of intra-flow. This new data editing process is described in the last chapter of this paper.

Introduction

La loi de modernisation de l'économie de 2008 donne une définition économique de l'entreprise¹. Afin de prendre en compte cette définition pour le calcul des statistiques structurelles d'entreprises, il a été décidé de prendre en compte la notion de groupe de société dans le système de production ESANE (Elaboration des Statistiques Annuelles des Entreprises). Si pour les plus grands groupes (qu'on appelle groupe de la cible 1), le profilage est réalisé manuellement, il n'est pas possible d'étendre cette méthode à l'ensemble des groupes restants (qu'on appelle groupes de la cible 2), pour des raisons évidentes de moyen. Ainsi l'enjeu pour le profilage automatique est de considérer les petits groupes comme des entreprises (et non plus les unités légales dont ils sont formés). Toutefois, comme les sources fiscales et les données d'enquêtes ne sont disponibles qu'au niveau des unités légales (UL), il a été décidé de distinguer les unités de collecte (qui restent les UL) des unités statistiques (entreprise). La valeur des caractéristiques « consolidées » au niveau de l'entreprise seront obtenues par une méthode, dite de bottom-up : le point de départ restant la somme des caractéristiques des UL. L'idée sous-jacente à cette consolidation est de retirer les chiffres d'affaires des unités légales du groupe qui est réalisé avec d'autres unités légales du groupe². Ce faisant, on va diminuer le chiffre d'affaires du groupe pour obtenir une estimation du chiffre d'affaires consolidé un peu meilleure que la somme brute des chiffres d'affaires des unités légales.

L'objectif de ce papier est de présenter la méthode bottom-up mise en place pour l'estimation du chiffre d'affaires, de confronter les résultats obtenus à des chiffres d'affaires consolidés par les groupes eux-mêmes et dont on dispose pour quelques groupes, puis de montrer comment est prise en compte la distinction entre unité de collecte et unité statistique lors de la phase de contrôle des données d'ESANE pour l'élaboration des statistiques structurelles d'entreprises..

1. Principes de consolidation et confrontation des résultats

L'algorithme se base sur plusieurs hypothèses fondamentales nécessaires pour estimer des flux intra-groupes:

- L'UL étant l'unité de collecte, on dispose de toutes les réponses des UL des contours des entreprises ou d'une estimation de leur réponse. Ceci a nécessité de revoir le plan de sondage de l'ESA[1].
- Quand il y a des unités qu'on détermine comme étant auxiliaires³ au sein d'un groupe, on suppose qu'elles n'accèdent pas au marché et **ne servent que le groupe** ;

¹ L'entreprise est la plus petite combinaison d'unités légales qui constitue une unité organisationnelle de production de biens et de services jouissant d'une certaine autonomie de décision, notamment pour l'affectation de ses ressources courantes.

²L'hypothèse est que ces échanges se font « hors marché » et qu'il n'ont donc pas à être pris en compte dans nos agrégats.

³ Nous verrons plus loin comment.

- Quand des unités seront considérées comme « intégrées » au sein d'une même entreprise, on fait l'hypothèse que l'unité en amont de la relation vende à l'unité en aval **au maximum** des capacités d'achat de l'unité aval et des capacités de production (ou ventes) de l'unité amont (c'est-à-dire **au minimum de ces deux capacités**);
- Quand on ne peut pas faire d'autres hypothèses que les 2 précédentes, et qu'ainsi on n'a aucune source d'information sur les flux d'échanges internes au sein des groupes, on fait l'hypothèse **que ces flux sont nuls**. En effet on n'a pas d'information sur la consolidation au sein de ces petits groupes qui n'ont d'ailleurs généralement pas d'obligation de produire des comptes consolidés. Dans ce cas où on aucune UL n'est considérée comme auxiliaire ou intégrée on ne peut pas estimer de flux potentiels, on se résout à estimer le compte consolidé de l'entreprise à la somme des comptes des UL de son contour.

1.1. Principes de consolidation

Étant donné le périmètre d'une entreprise, la constitution d'un compte consolidé s'effectue par étapes successives :

Étape 1 : on détermine le caractère auxiliaire, commercial ou productif des UL qui composent l'entreprise, à partir de l'Activité Principale Exercée (APE) de l'Unité Légale (UL). Une UL est considérée

- comme auxiliaire(A) : si son activité est orientée vers l'entreprise, son activité est donc considérée comme totalement intra-groupe . Une UL dans une entreprise est classée comme auxiliaire si son APE appartient à une liste d'activités possibles d'une part et si le poids en termes d'emploi au sein du groupe G de toutes les unités ayant cette APE est inférieur aux seuils suivants, pour les niveaux de nomenclatures suivants :
 - **10 %:** 47.78C, 49.41B, 64.30, 8559A, 62, 66, 68, 77
 - **20%:** 78.10, 69.1, 69.2, 64.2, 70.1, 71.1
 - pour les UL d'APE : 64.2, 70.1, 78.3 le caractère auxiliaire ne fait pas de doute et aucun seuil d'effectif n'est nécessaire.⁴
- comme commerciale(C) si elle pratique une activité de commerce (les deux premiers chiffres de l'APE sont compris entre 45 et 47)
- comme productive (P), toutes les UL non classées comme auxiliaires ou commerciales .

Étape 2 : Le but de cette étape est de calculer indépendamment le CA consolidé des 2 sous-groupes contenant l'ensemble des UL productives d'une part et commerciales d'autre part puis de consolider ces 2 sous-groupes entre eux.

Pour ce faire on utilise 3 listes d'activités intégrées.

- La première est une liste d'activités productives amont et aval. L'hypothèse faite est que la production de l'entreprise en amont sert de matière première⁵ pour la production de l'entreprise en aval. Dans ce cas on consolide au maximum le CA de l'UL en amont avec les achats de matière première de l'UL en aval.
- La deuxième est une liste d'activités commerciales de gros et de détail. L'hypothèse faite est que l'unité de commerce de détail s'approvisionnent chez l'UL de commerce de gros de la même entreprise. Dans ce cas on consolide au maximum le CA de l'UL de commerce de gros avec les achats de marchandise de l'UL du commerce de détail.

⁴Bien entendu si l'APE du groupe est auxiliaire, les UL du périmètre ne sont plus considérées comme telles et leur chiffre d'affaires n'est pas consolidé..

⁵Ces échanges à l'intérieur du groupe, hors marché, ne sont pas à prendre en compte sous peine de double compte.

- La troisième une liste d'activités productives et commerciales. L'hypothèse faite est que l'unité productive vende sa production à l'entreprise commerciale du groupe qui se charge de sa vente sur le marché. Dans ce cas on consolide au maximum le CA de la productive avec les achats de marchandises de la commerciale.

Ces différentes hypothèses conduisent à la définition d'une typologie possible de liens entre UL.

Tableau 1 : relations comptables entre les UL du contour

Type de lien	Par UL	Relations comptables entre les UL du groupe
A1	A et (P ou C)	Le chiffre d'affaires de l'auxiliaire est comptabilisé dans le poste « autres charges d'exploitations » de la productive (ou la commerciale)
A2	A et (P ou C)	L'APE de l'auxiliaire quelconque sauf 64.2, 68, 77
A3	A et (P ou C)	APE de l'auxiliaire=64.2 (holding)
A4	A et (P ou C)	APE de l'auxiliaire =68 (activités immobilières) ou 77 (location ou location-bail)
B1	P et C	APE de la commerciale =46.1 (intermédiaires du commerce de gros)
B2	P et C	Le chiffre d'affaires de la productive inférieur aux achats de marchandises de la commerciale
B3	P et C	Le chiffre d'affaires de la productive supérieur aux achats de marchandises de la commerciale
C1	P et P	Les ventes de biens de la productive en amont sont inférieures au poste « achats de matières premières et autres approvisionnement » de la productive en aval
C2	P et P	Les ventes de biens de la productive en amont sont supérieures au poste « achats de matières premières et autres approvisionnement » de la productive en aval
D1	C et C	Les ventes de marchandises du commerce de gros sont inférieures aux achats de marchandises du commerce de détail
D2	C et C	Les ventes de marchandises du commerce de gros sont supérieures aux achats de marchandises du commerce de détail
E1		Les 2 UL ont le même APE
E2		Consolidation simple : aucun des critères précédents n'est retenu.

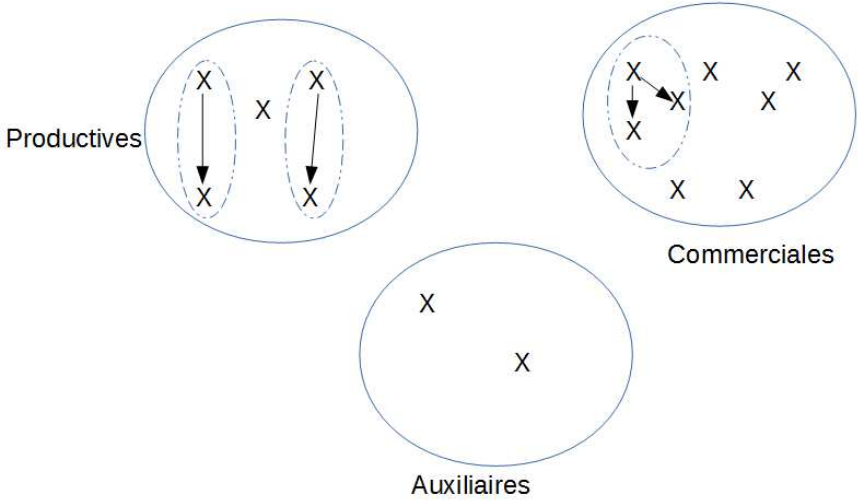
Effet de la consolidation sur le chiffre d'affaires selon le type de groupe :

Tableau 2 : Composantes à retirer du chiffre d'affaires

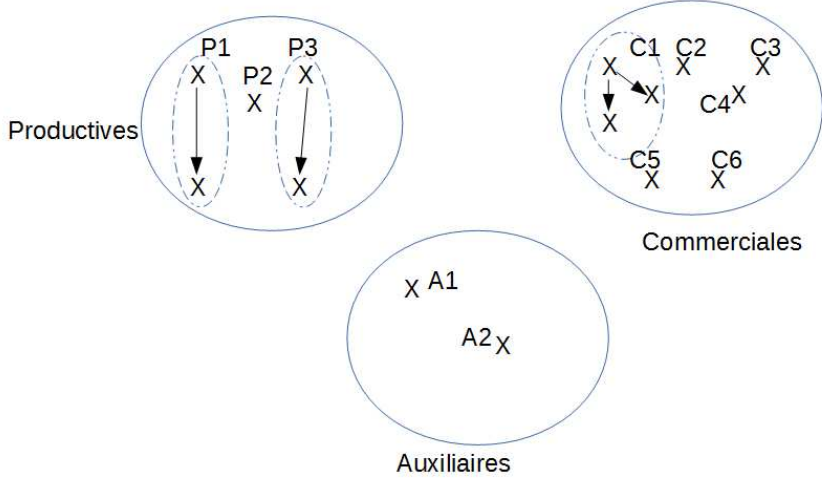
Type de groupe	Par UL	Composantes à retirer du chiffre d'affaires de l'entreprise profilée
A1	A et (P ou C)	on retire le chiffre d'affaires de l'auxiliaire
A2	A et (P ou C)	on retire le chiffre d'affaires de l'auxiliaire
A3	A et (P ou C)	on retire le chiffre d'affaires de l'auxiliaire
A4	A et (P ou C)	on retire le chiffre d'affaires de l'auxiliaire
B1	P et C	on retire la composante « Ventes de services » de la commerciale
B2	P et C	on retire les composantes « Ventes de biens » et « Ventes de marchandises » de la productive
B3	P et C	on retire la composante « Achats de marchandises » de la commerciale
C1	P et P	on retire la composante « Ventes de biens » de la productive en amont
C2	P et P	on retire la composante « Achats de matières premières et autres approvisionnements » de la productive en aval
D1	C et C	on retire la composante « Ventes de marchandises » du commerce de gros
D2	C et C	on retire la composante « Achats de marchandises » du commerce de détail

La liste d'activités intégrées pour les activités productives peut conduire à ce qu'une même UL soit en aval d'une seconde, mais en amont d'une troisième. Par exemple 0220 (exploitation forestière) vend à 1610 (sciage et rabotage du bois), qui vend à 1621 (fabrication de placage et de panneaux de bois). Dans ce cas-là on gardera un seul et unique flux. Ce sera celui dont l'UL a les achats de matières premières les plus élevés. Cela permettra de conserver un maximum d'effet de consolidation, mais de ne pas rajouter de la complexité pour les gestionnaires qui auront en charge de vérifier les liens de consolidation. En 2014, seulement 17 entreprises profilées rencontrent ce cas.

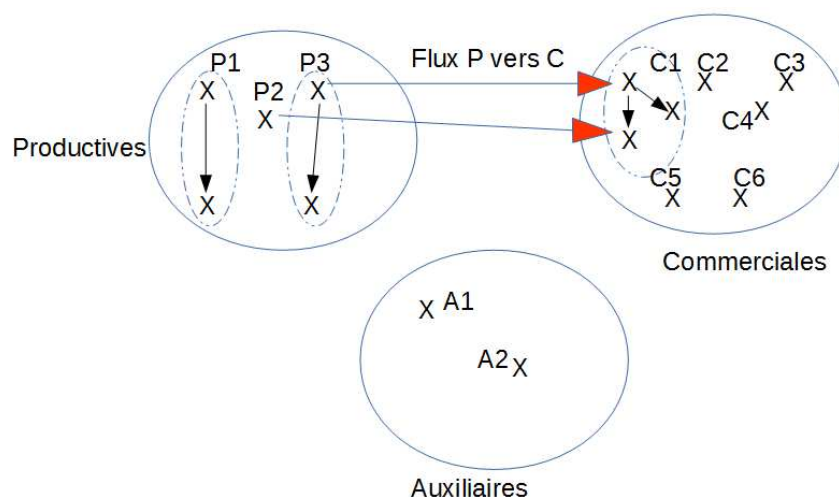
On procède de la même façon pour les UL commerciales, avec les UL de commerce de gros qui vendent aux UL de commerce de détail. On obtient ainsi des binômes Pamont vers Paval, et Cgros vers Cdétail, que l'on va consolider.



On obtient ainsi dans chaque ensemble d'UL (P, C ou A) des UL qui sont binômées et certaines qui ne sont pas binômées. On peut représenter cette action ainsi : pour P1, P3 et C1 ces "sous-groupes" sont consolidés.



Étape 3 : de la même façon, à partir de la liste d'APE de productives qui vendent à des commerciales, on va chercher à repérer les binômes Productives – Commerciales au sein de l'entreprise. Si au moins une UL d'un sous-groupe productive est potentiellement en lien avec une UL du sous-groupe commerciale selon la liste 3 évoquée ci-dessus, alors on consolide l'ensemble des 2 sous- groupes déjà consolidés (flèches rouges). On va donc ici consolider l'ensemble (P2+P3) avec l'ensemble C1.



Étape 4 : enfin on consolide le CA du sous groupe des UL auxiliaires avec les sous-groupes consolidés dans les étapes précédentes qui regroupent toutes les UL commerciales et productives. Comme ces UL auxiliaires ne sont considérées orientées vers le marché, mais uniquement vers l'entreprise, on considère la totalité de leur chiffre d'affaires comme un flux intra, il est donc supprimé.

Exemple de consolidation d'une entreprise :

Le tableau suivant présente les variables d'achats et de ventes pour les 15 UL du contour d'une entreprise :

	type_UL	Sous-groupe	Achats de marchandises	Achats de matières premières	Ventes de biens	Ventes de marchandises	Ventes de services	CA
1	P	P1	0	110	136	0	0	136
2	P	P1	0	130	157	0	0	157
3	P	P2	0	15	17	0	0	17
4	P	P3	0	180	215	0	0	215
5	P	P3	0	190	238	0	0	238
6	C	C1	285	0	0	295	0	295
7	C	C1	170	0	0	198	0	198
8	C	C1	50	0	0	85	0	85
9	C	C2	180	0	0	207	0	207
10	C	C3	120	0	0	155	0	155
11	C	C4	170	0	0	201	0	201
12	C	C5	50	0	0	81	0	81
13	C	C6	10	0	0	16	0	16
14	A	A1	0	0	0	0	100	100
15	A	A2	0	0	0	0	200	200

Le type_UL est une variable qui distingue chaque UL en productives P, commerciales C ou auxiliaires A (qui correspond à l'étape 1 décrite au 1.1). Les sous-groupes qui ont pour identifiant P1, P2 ... reprennent les exemples des sous-groupes présentés plus haut. Chaque sous-groupe pour lequel on va effectuer une consolidation est repéré par une couleur.

Le chiffre d'affaires de la somme des UL est 2 301.

La consolidation débute avec les "sous-groupes" P1, P3 et C1 (qui correspond à l'étape 2 décrite au 1.1) :

dans P1 et P3, on retire les achats de matières premières des UL en aval car il est inférieur au CA de l'UL amont. Le CA (tout comme les ventes de biens) est donc amputé : pour P1 de -130, et pour P3 de -190. Dans C1, on retire les achats de marchandises des UL de commerce de détail car ils sont inférieurs aux ventes de l'UL de commerce de gros qui vaut 295. Le CA est donc amputé de $-170-50=-220$.

La consolidation continue avec le flux P vers C (qui correspond à l'étape 3 décrite au 1.1) : c'est le lien de la flèche rouge entre (P2+P3) et C1. On retire les ventes de marchandises et de biens des UL productives, qui valent : $17+215+238 - 190=280$ pare qu'elles sont inférieure aux achats de C1 qui valent $285 + 170 + 50 - 220 = 285$. Le terme -190 correspond à la partie consolidée des ventes de biens dans le "sous-groupe" P3. Le CA des UL restantes non binômées est sommé simplement.

Enfin on effectue la consolidation avec les UL auxiliaires. On ampute leur CA : -300.

Au total, pour avoir le CA de l'entreprise profilée, le CA sommé simplement est amputé de $-130-190-220-280-300=-1\ 120$.

Le chiffre d'affaires de l'entreprise profilée est donc $2\ 301-1\ 120 = 1\ 181$.

2. Effets de la consolidation sur le chiffre d'affaires

Le tableau suivant récapitule les flux intra-groupe de chiffre d'affaires par type de flux : auxiliaires, productives vers commerciales, productives vers productives et commerciales vers commerciales.

Tableau 3 : effet de la consolidation sur l'ensemble des groupes de la cible 2 par type de flux

	intra CA
A	-58 198 624
P+C	-86 149 509
P+P	-8 948 884
C+C	-10 493 127
Total	-163 790 144

La moitié du flux intra total de -164 milliards d'euros de chiffre d'affaires est expliquée par les flux intra entre productives et commerciales.

Le tableau suivant présente les effets de la consolidation par grands secteurs. On compare le CA des entreprises profilées au CA obtenu en faisant la somme des UL qui composent l'EP.

Tableau 4 : effet de la consolidation en A10 par APE de l'entreprise profilée

	flux intra
Industrie	-8,8%
Construction	-5,0%
Commerce transport hébergement restauration	-6,6%
Information et communication	-3,9%
Activités immobilières	-3,4%
Activités spécialisées, scientifiques et techniques et services administratifs et de soutien	-6,4%
Administration publique, enseignement, santé humaine et action sociale	-2,9%
Autres activités de services	-1,5%
Total	-7,0%

En 2013 pour la cible 1, l'effet de la consolidation se traduisait par une réduction du CA de 13% des UL qui constituent les EP. L'effet de la consolidation est le plus important pour les EP industrielles. C'est en effet dans ce secteur qu'un grand nombre d'EP présentent une consolidation importante. Cela s'explique par le fait qu'une EP industrielle peut éventuellement bénéficier de plus de flux intra que toutes les autres EP. En effet elle peut présenter un flux Pamont vers Paval, Cgros vers Cdétail, P vers C, et de la consolidation des UL auxiliaires. À l'inverse une EP commerciale ne bénéficie généralement que des flux Cgros vers Cdétail, et de la consolidation des UL auxiliaires.

Le tableau suivant indique le nombre d'EP par effet de consolidation par APE de l'EP.

Tableau 5 : répartition du nombre d'EP selon l'effet de la consolidation par secteur de l'EP

APE de l'entreprise	moins de 10 %	entre 20% et 10%	entre 30% et 20%	entre 40% et 30%	entre 50% et 40%	entre 60% et 50%	entre 70% et 60%	entre 80% et 70%	entre 90% et 80%	entre 100% et 90%	Total
INDUSTRIES EXTRACTIVES	165	13	8	6						1	193
INDUSTRIE MANUFACTURIÈRE	13597	1468	406	171	45	8	3	2	1	2	15703
PRODUCTION ET DISTRIBUTION D'ÉLECTRICITÉ, DE GAZ, DE VAPEUR ET D'AIR CONDITIONNÉ	348	21	3		2						374
PRODUCTION ET DISTRIBUTION D'EAU ; ASSAINISSEMENT, GESTION DES DÉCHETS ET DÉPOLLUTION	441	26	6	1	2					1	477
CONSTRUCTION	9514	695	102	27	8	1	3	3	1	6	10360
COMMERCE ; RÉPARATION D'AUTOMOBILES ET DE MOTOCYCLES	23468	1147	366	181	46	21	10	9	10	13	25271
TRANSPORTS ET ENTREPOSAGE	2640	194	43	9	1	2					2889
HÉBERGEMENT ET RESTAURATION	4470	485	96	22	8	1			1	2	5085
INFORMATION ET COMMUNICATION	3339	284	91	21	7	1	4				3747
ACTIVITÉS IMMOBILIÈRES	2451	164	69	19	5	1					2709
ACTIVITÉS SPÉCIALISÉES, SCIENTIFIQUES ET TECHNIQUES	5506	918	506	105	27	7		5		8	7082
ACTIVITÉS DE SERVICES ADMINISTRATIFS ET DE SOUTIEN	3147	293	66	15	6			1	2		3530
ENSEIGNEMENT	513	63	28	3	1			1			609
SANTÉ HUMAINE ET ACTION SOCIALE	1147	71	16	3			1				1238
ARTS, SPECTACLES ET ACTIVITÉS RÉCRÉATIVES	578	41	18	4	1						642
AUTRES ACTIVITÉS DE SERVICES	663	75	32	6	1			1			778
Total	71987	5958	1856	593	160	42	21	22	15	33	80687

Lecture : parmi les 160 EP qui ont un effet de consolidation sur le CA entre -40 et -50%, 45 sont dans le secteur manufacturier C.

La majorité des entreprises ont une consolidation inférieure à 10 %. Seulement 11 % des entreprises ont une consolidation supérieure à 10 %.

Le tableau ci-dessous présente pour 2014 l'effet de la prise en compte de l'APE sur le secteur soit de l'UL, soit de l'EP, ainsi que l'effet de la consolidation.

- Ainsi, la première colonne est le CA calculé par APE de l'UL,
- la deuxième colonne représente le CA calculé au niveau de l'APE de l'EP,
- la troisième colonne représente le CA consolidé calculé au niveau de l'APE de l'EP.

Tableau 6 : répartition du CA (en K€) pour les UL appartenant à des groupes de la cible 2

	CA par APE de l'UL	%	CA par APE du groupe	%	CA consolidé par APE du groupe	%
Agriculture, sylviculture et pêche	250 736	0	0	0	0	0
Industrie manufacturière, industries extractives et autres	697 306 497	30	808 623 964	35	737 262 256	34
Construction	110 169 044	5	113 514 974	5	107 866 650	5
Commerce de gros et de détail, transports, hébergement et restauration	1 122 107 257	48	1 049 373 019	45	979 719 363	45
Information et communication	98 845 895	4	94 419 472	4	90 703 858	4
Activités financières et d'assurance	20 169 711	1	0	0	0	0
Activités immobilières	35 846 163	2	36 213 276	2	34 994 444	2
Activités spécialisées, scientifiques et techniques et activités de services administratifs et de soutien	190 986 137	8	169 480 843	7	158 613 436	7
Administration publique, enseignement, santé humaine et action sociale	27 371 075	1	28 775 444	1	27 948 775	1
Autres activités de services	31 349 041	1	34 000 564	1	33 502 630	2
Total	2 334 401 556	100	2 334 401 556	100	2 170 611 411	100

Lecture : Les UL du secteur de l'industrie appartenant à une entreprise profilée réalisent un chiffre d'affaires de 697 306 497 K€. Les entreprises profilées de la cible 2 du secteur de l'industrie manufacturière réalisent un chiffre d'affaires non consolidée de 808 623 964 K€ et un chiffre d'affaires consolidé par l'algorithme de 737 262 256 K€.

On constate donc que la déformation de la répartition sectorielle entre la vision en UL et en entreprise provient de la conjugaison de deux effets :

- un effet de réallocation (écart entre les colonnes 2 et 4 du tableau). Cet effet provient du fait que toutes les UL d'une entreprise, quel que soit leur secteur sont comptabilisé dans le secteur de l'entreprise dans la vision en entreprise. En effet, Les entreprises industrielles ou de construction constituées en groupe comprennent de nombreuses sociétés dans leur coeur de métier. Cependant, elles ont aussi souvent créé des sociétés distinctes ayant une activité principale pour réaliser leurs fonctions commerciales en France ou à l'exportation, de même que leurs fonctions supports (holding, siège social, transport, immobilier, recherche, etc.). Ainsi, le secteur industriel voit le poids relatif de son chiffre d'affaires s'accroître 30 à 35 % ors du passage d'une approche en unités légales à une approche en entreprises :
- l'effet de la consolidation (écart entre les colonnes 4 et 6 du tableau). Cet effet est de 7 % au global comme déjà évoqué ci-dessus.

On constate via ce tableau que c'est principalement l'effet de réallocation qui explique l'écart de vision entre les deux approches. Ce résultat est en phase avec ce qui a été observé sur les entreprises profilées manuellement.

3. Confrontation des données

3.1 Présentation des données banque de France

La Banque de France dispose de comptes consolidés de groupes. En effet certains groupes ont l'obligation légale de fournir des comptes consolidés lorsqu'ils dépassent certains seuils de chiffre

d'affaires ou d'effectifs, ou lorsqu'ils souhaitent être coté sur un marché. On a ainsi cherché à mesurer la qualité de notre algorithme au regard du « vrai » chiffre d'affaires consolidé par le groupe. Pour l'évaluation de l'algorithme ce sont les données de l'année 2013 qui sont utilisées. L'idée de départ étant d'éviter autant que faire se peut que l'algorithme conduise à une sur-consolidation, on va surtout vérifier que dans la majeure partie des cas notre CA consolidé par notre algorithme est supérieur au CA consolidé par le groupe lui-même. En effet on sait actuellement que l'on surestime le CA total en sommant simplement les CA de chaque UL du contour (soit en considérant les flux intra groupe nuls). Notre hypothèse de travail est de réduire cette erreur. Toutefois on se donne également comme contrainte de ne pas réduire l'activité économique réelle et donc d'essayer autant que faire se peut de ne jamais trop consolider.

Le fichier de la Banque de France est constitué de 4 976 observations a priori pour l'année 2013. Néanmoins en filtrant sur l'année 2013, on obtient 4 070 groupes. Les SIREN de tête de groupe sont exclusivement des SIREN français. Lorsque la durée de l'exercice est différente de 12 mois, on cale le chiffre d'affaires consolidé sur 12 mois.

On étudie la dispersion de l'écart entre le chiffre d'affaires consolidé Insee et le chiffre d'affaires consolidé Banque de France, en points de pourcentage.

Le tableau suivant récapitule le nombre de groupes

- pour l'Insee : construits à partir de la base des groupes de l'Insee (LIFI), qui ont une liasse fiscale non nulle hors groupes financiers et agricoles
- pour la Banque de France : filtrés sur l'année de centralisation

	Insee	Banque de France
Nombre de groupes	81 100	4 070

3.2 Les groupes retenues pour la comparaison

On s'intéresse aux groupes qui sont simultanément présents dans les 2 bases. L'appariement est réalisé sur le SIREN de la tête de groupe. 2 789 groupes sont présents dans les 2 bases (dont 695 qui ont au maximum 2 activités soit 25%), 1 281 groupes sont présents dans la base Banque de France, mais absents de la base Insee (31% des groupes Banque de France).



Si on apparie les 1 281 SIREN (qui sont pour la Banque de France des SIREN de tête de groupe, mais pour LIFI des SIREN d'UL) avec la table de 78 311 groupes Insee, soit 240 097 UL : 203 SIREN sont présents.

Les SIREN de tête de groupe correspondants à ces 203 SIREN d'UL vont pouvoir être ajoutés aux 2 789 SIREN de tête de groupe déjà appariés. On aurait ainsi $2\,789 + 203 = 2\,992$ **groupes présents à la fois dans la table Insee et la table Banque de France** (soit 91% des groupes de la table Banque de France filtrée).

Le périmètre en UL des groupes de la Banque de France n'est pas défini dans les données. Ainsi pour s'assurer que le contour du groupe consolidé de la Banque de France est le même que celui dont on dispose et ainsi comparer les CA sur un contour identique, on va ajouter un filtre sur l'effectif en équivalent temps plein (EQTP). L'effectif français du groupe de Insee doit être proche de l'EQTP Banque de France (+/- 15%), et le chiffre d'affaires consolidé de la Banque de France doit être inférieur au CA de la somme des UL du contour connu de LIFI⁶ (ce qui correspond dans le tableau au "filtre contour"). **1 011 groupes ont ainsi des contours proches.**

3.3 Les écarts de consolidation observés

Le tableau suivant présente ces résultats :

Tableau 7 : répartition des groupes selon la différence de consolidation entre la consolidation automatique et celle fournie par la Banque de France

Diff en points de %	Ensemble des groupes	
	sans filtre	avec le filtre contour
Nb de groupes	2 292	1 011
Max	100,0	100,0
99%	46,1	43,7
95%	25,3	28,5
90%	16,4	20,3
75% Q3	6,5	10,6
50%	0,2	3,5
25% Q1	-13,1	0,2
10%	-64,4	-2,5
5%	-145,1	-8,5
1%	-1 700,9	-21,1
Min	-128 947,0	-39,0

Lecture : avec le filtre contour, 1 011 groupes ont été appariés entre la base Insee et la base Banque de France, et dans 1 % de ces groupes, l'effet de la consolidation Insee est inférieure de 43,7 points à la consolidation de la Banque de France.

La répartition des 2 292 groupes sont des ETI à 56% et des PME à 44%. Concernant les 1 011 groupes, la répartition est identique.

Lorsqu'on compare la différence de consolidation sur uniquement les groupes qui ont un contour proche, les différences sont faibles. Ainsi si on prend tous les groupes avec le filtre contour, on consolide trop d'au moins 8,5 points de plus que la Banque de France dans seulement 5 % des cas. Donc en généralisant grossièrement, on aurait seulement 5% des 81 100groupes qui auraient une consolidation trop importante de plus de 8,5 points, soit 4 000 groupes.

⁶En effet, par définition le chiffre d'affaires consolidé est inférieur au CA des UL du périmètre. Si ce n'est pas le cas cela signifie que le périmètre n'est pas le bon.

La partie grisée indique que 80% des groupes ont très peu de différence de consolidation du CA (environ entre 0% et 20%, ce qui signifie que l'algorithme automatique a bien ciblé les flux internes ou n'a pas trop consolidé par rapport à la réalité). Ainsi notre consolidation est jugée plutôt prudente ce qui répond à l'une des contraintes initiales qu'on s'était fixée.

3.3.1 Différences de consolidation par secteurs

Néanmoins on va s'intéresser aux groupes pour lesquels nos principes de consolidation automatique conduisent à une surestimation ou une sous-estimation de la consolidation. Par exemple si l'effet de consolidation Insee est de -10%, et celui de la Banque de France est de -8%, dans ce cas-là on aura un écart de points de -2, et l'algorithme aura trop consolidé.

Le tableau suivant indique dans l'avant-dernière colonne la répartition par secteurs du nombre de groupes qui ont une consolidation d'au moins 10 points de trop par rapport à la consolidation de la Banque de France.

Tableau 8 : répartition du nombre de groupes dont la consolidation automatique est trop forte ou trop faible par rapport à la consolidation Banque de France

a21	groupes avec filtre contour		dont groupes conso forte		dont groupes conso faible	
INDUSTRIES EXTRACTIVES	4	0 %			1	1 %
INDUSTRIE MANUFACTURIÈRE	255	25 %	10	24 %	27	26 %
PRODUCTION ET DISTRIBUTION D'ÉLECTRICITÉ, DE GAZ, DE VAPEUR ET D'AIR CONDITIONNÉ	2	0 %			1	1 %
PRODUCTION ET DISTRIBUTION D'EAU ; ASSAINISSEMENT, GESTION DES DÉCHETS ET DÉPOLLUTION	8	1 %			3	3 %
CONSTRUCTION	80	8 %	1	2 %	9	9 %
COMMERCE ; RÉPARATION D'AUTOMOBILES ET DE MOTOCYCLES	405	40 %	16	39 %	29	28 %
TRANSPORTS ET ENTREPOSAGE	63	6 %	1	2 %	7	7 %
HÉBERGEMENT ET RESTAURATION	25	2 %				
INFORMATION ET COMMUNICATION	32	3 %			4	4 %
ACTIVITÉS IMMOBILIÈRES	14	1 %	1	2 %	3	3 %
ACTIVITÉS SPÉCIALISÉES, SCIENTIFIQUES ET TECHNIQUES	58	6 %	10	24 %	12	12 %
ACTIVITÉS DE SERVICES ADMINISTRATIFS ET DE SOUTIEN	32	3 %	1	2 %	5	5 %
ENSEIGNEMENT	1	0 %				
SANTÉ HUMAINE ET ACTION SOCIALE	22	2 %	1	2 %		
ARTS, SPECTACLES ET ACTIVITÉS RÉCRÉATIVES	6	1 %			1	1 %
AUTRES ACTIVITÉS DE SERVICES	4	0 %				
TOTAL	1011	100 %	41	100 %	102	100 %

Lecture : 41 groupes ont une consolidation d'au moins 10 points de trop par rapport à la Banque de France. Ce sont surtout des groupes industriels et commerciaux qui les composent. 102 groupes ont une consolidation d'au moins 20 points trop faible par rapport à la Banque de France.

Ces groupes sont majoritairement des groupes industriels (BE) ou commerciaux (GZ), et ceux des activités spécialisées, scientifiques et techniques (M) ce qui est représentatif de la composition de l'ensemble des groupes de la cible 2.

Le tableau suivant présente la répartition des 1 011 groupes selon différents chiffres d'affaires. **Le cumul du chiffre d'affaires consolidé par l'algorithme est toujours supérieur au cumul du chiffre d'affaires de la Banque de France, par grand secteur.** On retrouve les effets de la consolidation les plus importants sur les groupes industriels, commerciaux, et ceux des activités spécialisées, scientifiques et techniques.

Tableau 9 : répartition en K€ du CA du cumul des liasses, du CA consolidé par l'algorithmme et du CA fourni par la Banque de France pour les 1 011 groupes

	Sans consolidation			Consolidation Insee			Consolidation BdF	
	Cumul des liasses	nb de groupes	poids en CA dans la cible 2	Cumul des CA consolidés	diff en % par rapport au cumul des liasses	nb de groupes avec une conso > 10 points au CA BdF	Cumul des CA consolidés	diff en % par rapport au cumul des liasses
INDUSTRIES EXTRACTIVES	337 333	4	10%	302 520	-10%		299 998	-11%
INDUSTRIE MANUFACTURIERE	25 591 977	255	3%	22 873 057	-11%	10	21 416 494	-16%
PRODUCTION ET DISTRIBUTION D'ÉLECTRICITÉ, DE GAZ, DE VAPEUR ET D'AIR CONDITIONNÉ	1 219 210	2	2%	1 217 325	0%		922 741	-24%
PRODUCTION ET DISTRIBUTION D'EAU ; ASSAINISSEMENT, GESTION DES DÉCHETS ET DÉPOLLUTION	1 225 333	8	3%	1 213 689	-1%		1 037 815	-15%
CONSTRUCTION	7 359 957	80	6%	6 984 406	-5%	1	6 457 204	-12%
COMMERCE ; RÉPARATION D'AUTOMOBILES ET DE MOTOCYCLES	75 464 554	405	10%	70 637 295	-6%	16	64 248 138	-15%
TRANSPORTS ET ENTREPOSAGE	50 237 736	63	31%	48 191 687	-4%	1	32 432 652	-35%
HÉBERGEMENT ET RESTAURATION	1 850 695	25	5%	1 787 873	-3%		1 646 237	-11%
INFORMATION ET COMMUNICATION	2 524 491	32	3%	2 330 779	-8%		2 068 085	-18%
ACTIVITÉS IMMOBILIÈRES	1 235 600	14	4%	1 166 050	-6%	1	1 023 040	-17%
ACTIVITÉS SPÉCIALISÉES, SCIENTIFIQUES ET TECHNIQUES	4 070 303	58	5%	3 728 609	-8%	10	3 472 606	-15%
ACTIVITÉS DE SERVICES ADMINISTRATIFS ET DE SOUTIEN	2 700 248	32	3%	2 666 555	-1%	1	2 474 543	-8%
ENSEIGNEMENT	48 139	1	1%	48 139	0%		47 591	-1%
SANTÉ HUMAINE ET ACTION SOCIALE	1 469 681	22	6%	1 408 304	-4%	1	1 382 889	-6%
ARTS, SPECTACLES ET ACTIVITÉS RÉCRÉATIVES	15 494 030	6	66%	15 420 856	0%		13 998 923	-10%
AUTRES ACTIVITÉS DE SERVICES	221 377	4	3%	219 283	-1%		205 693	-7%
Total	191 050 663	1 011	8%	180 196 427	-6%	41	153 134 649	-20%

Lecture : Dans le secteur manufacturier C, le chiffre d'affaires sommé simplement s'élève à 25,592 milliards d'euros pour 255 groupes qui pèsent en terme de CA 3 % dans la cible 2. Le cumul des CA consolidés par l'algorithmme s'élève à 22,873 milliards d'euros, soit 11 % de moins que le CA sommé simplement. 10 groupes possèdent un effet de consolidation sur le CA trop fort d'au moins 10 points par rapport à la Banque de France. Le cumul des CA consolidés par la Banque de France s'élève à 21,416 milliards d'euros. C'est 16 % de moins que le CA sommé simplement.

3.3.2 Entreprises dont la consolidation par algorithmme est trop importante

Si on considère les 10 plus gros écarts à la consolidation Banque de France, on remarque que ce sont surtout des groupes dans lesquels on a fait une consolidation de type P → C. La consolidation fournie par la Banque de France aurait ainsi tendance à dire qu'une partie des flux de P ne vont pas vers C, et donc qu'il ne faut pas consolider entièrement ces flux. En effet après l'examen des 41 groupes dont la consolidation est trop forte et de l'appariement de type P → C, ils présentent tous des flux intra-groupes probables (de par les APE des UL qui constituent l'EP) et il n'existe qu'un seul flux unique de ce type par EP.

3.3.3 Entreprises dont la consolidation par algorithmme n'est pas assez importante

La dernière colonne du tableau 8 présente les groupes pour lesquels la consolidation automatique n'est pas suffisante par rapport à la consolidation fournie par la Banque de France. Elle décrit la répartition par secteurs du nombre de groupes qui ont une consolidation sur le chiffre d'affaires d'au moins 20 points trop faible par rapport à la consolidation Banque de France. Comme pour les groupes dont la consolidation automatique est trop forte, ces groupes sont majoritairement des groupes industriels (BE) ou commerciaux (GZ), et ceux des activités spécialisées, scientifiques et techniques (M) ce qui est également représentatif de la composition de l'ensemble des groupes de la cible 2.

Parmi ces 102 groupes, seulement 31 n'ont aucune consolidation. La majorité des autres groupes n'ont qu'un seul type de consolidation qui est dû aux UL auxiliaires.

817 groupes ont un chiffre d'affaires consolidé supérieur à celui de la Banque de France. Ils sont constitués à 53 % d'ETI et à 47 % de PME. L'extrapolation de ce flux intra des 817 groupes représentent donc un montant de chiffre d'affaires qu'on aurait pu ajouter aux 164 milliards d'euros de flux intra déjà estimés par notre méthode (cf. 2.). Pour mesurer l'effet de ce manque sur l'ensemble de la cible 2, on a donc essayé d'extrapoler le montant de ce flux intra. On a construit des classes de chiffre d'affaires sommé simplement pour l'ensemble des 81 100 EP de la cible 2. Sur ces classes on a calculé la médiane des écarts entre le chiffre d'affaires consolidé par algorithme et celui de la Banque de France parmi les 817 groupes (groupes dont le CA consolidé par algorithme est supérieur au CA consolidé de la Banque de France). On a pu ensuite appliquer cet écart aux chiffres d'affaires consolidés de chaque groupe dans chaque classe pour les 81 100 groupes de la cible 2 (cf. encadré 1).

Le total extrapolé sur la cible 2 de l'ensemble des 817 groupes dont on a pas assez consolidé le CA s'élève à 165 milliards d'euros.

3.3.4 Extrapolation des écarts observés à l'ensemble des groupes

Tableau 10 : effet de l'extrapolation sur le manque à consolider parmi les 817 groupes

Classe de CA en K€	Nb de groupes cible 2	Médiane des diff en % parmi les 817 groupes	Ecart
entre 0 et 25 000	71 618	6,9	22 390 145
25 000 et 40 000	3 247	5,5	5 480 496
40 000 et 60 000	1 977	4,4	4 095 094
60 000 et 85 000	1 152	5,9	4 628 232
85 000 et 120 000	848	5,1	4 185 716
> 120 000	2 258	8,4	124 552 430
Total	81 100		165 332 113

Lecture : dans la classe des groupes qui ont un CA compris entre 0 et 25 millions d'euros : cette classe est constituée de 71 618 groupes, et 50 % des groupes appartenant à cette classe parmi les 817 groupes ont une différence de CA entre celui fourni par l'algorithme et celui fourni par la Banque de France inférieur à 6,9 %.

Le même exercice réalisé sur l'ensemble des 1 011 groupes indique que le total extrapolé de l'ensemble des 1 011 groupes, soit les groupes dont on a fait trop de consolidation et ceux dont on n'en a pas fait assez, représente le manque de consolidation pour l'ensemble de la cible 2, soit **101 milliards d'euros supplémentaires aux 164 milliards d'euros précédemment calculés qui pourraient potentiellement compléter la consolidation du CA de l'ensemble des groupes de la cible 2.**

Tableau 11 : effet de l'extrapolation sur le manque à consolider parmi les 1 011 groupes

Classe de CA en K€	Nb de groupes cible 2	Médiane des diff en % parmi les 1 011 groupes	Ecart
entre 0 et 25 000	71 618	4,7	15 392 078
25 000 et 40 000	3 247	3,6	3 517 998
40 000 et 60 000	1 977	2,6	2 434 939
60 000 et 85 000	1 152	4,0	3 125 723
85 000 et 120 000	848	3,3	2 743 051
> 120 000	2 258	5,0	73 800 704
Total	81 100		101 014 494

Lecture : dans la classe des groupes qui ont un CA compris entre 0 et 25 millions d'euros : cette classe est constituée de 71 618 groupes, et 50 % des groupes appartenant à cette classe parmi les 1 011 groupes ont une différence de CA entre celui fourni par l'algorithme et celui fourni par la Banque de France inférieur à 4,7 %.

Encadré 1 :

Le calcul de l'écart total a été fait de la façon suivante :

$$\sum_{\text{groupe } i=1}^{81\ 100} \sum_{\text{classe } j} CA \text{ consolidé } pop_i \cdot \text{Médiane des écarts échantillon}_j \cdot \mathbf{1}_{i \in j}$$

avec $\mathbf{1} = \begin{cases} 1 & \text{si groupe } i \in \text{la classe } j \\ 0 & \text{sinon} \end{cases}$

Le fait d'appliquer la médiane des écarts dans l'échantillon au CA consolidé de chaque groupe dans chaque classe est justifié par le fait que dans chaque classe, les médianes des CA consolidés dans la population et dans l'échantillon sont proches. Ainsi, l'échantillon des 1 011 groupes est jugé représentatif de la population de la cible 2, comme l'indique le tableau suivant :

Classe de CA en K€	Nb de groupes cible 2	Médiane du CA conso cible 2	Nb de groupes échantillon	Médiane du CA conso échantillon
entre 0 et 25 000	71 618	2 523	119	14 632
25 000 et 40 000	3 247	29 987	177	31 567
40 000 et 600 000	1 977	46 832	208	46 481
60 000 et 85 000	1 152	68 003	154	68 293
85 000 et 120 000	848	97 061	116	94 266
> 120 000	2 258	238 680	237	192 047
	81 100		1 011	

La médiane du chiffre d'affaires consolidé de la cible 2 dans la classe « entre 0 et 25 000 » est inférieure à celle de l'échantillon. Cela conduit à une légère surestimation de notre manque de chiffre d'affaires à consolider.

Ce calcul n'est qu'une estimation fruste de la différence entre les 2 chiffres d'affaires mais indique bien un manque de consolidation a priori important.

Nos hypothèses sont donc peut être un peu trop prudentes mais elles ont le mérite de s'expliquer et il est difficile de faire mieux avec les informations en notre possession.

L'idée est donc d'essayer d'améliorer l'algorithme en collectant par le biais des enquêtes structurelles des données sur la part des échanges intra-groupes des UL.

4. Evolution de la chaîne de contrôle et de traitement des données

Les données publiées par ESANE seront calculées à partir des données des entreprises. Ainsi l'unité à prendre en compte pour la chaîne de contrôle d'ESANE est l'entreprise. Le travail des gestionnaires sera donc de valider les données au niveau « entreprise ». Mais les données collectées étant au

niveau UL, en cas de doutes sur une valeur, le gestionnaire devra contacter une UL directement et non l'entreprise. Il a donc fallu faire évoluer la chaîne de contrôle ESANE afin de prendre en compte cette distinction entre unité de collecte et d'intérêt.

4.1 Les grands principes

Toutes les UL non répondantes d'une entreprise seront extrapolées. Ainsi, chaque soir une consolidation permettra de disposer de données pour toutes les entreprises.

Pour une entreprise profilée les macro-contrôles ne seront calculés :

- que lorsque les UL répondantes représenteront au moins 70 % du chiffre d'affaires de l'entreprise
- et systématiquement à compter de début novembre, afin que toutes les entreprises puissent avoir été potentiellement contrôlées. En effet, le contrôle d'entreprises profilées incomplètes, même si les réponses des UL non répondantes sont systématiquement imputées, n'a pas vraiment de sens. Toute nouvelle réponse d'une UL du périmètre pourrait en effet conduire à une modification importante des données consolidées et donc à reconstruire une entreprise déjà analysée. En fin de campagne toutefois, les macro-contrôles seront réalisés sur l'ensemble des entreprises même si le poids des liasses imputées est supérieur à 30 %, ceci afin de s'assurer que tous les questionnaires seront contrôlés.

4.2 Le travail des gestionnaires

Une fois ce seuil dépassé ou après début novembre, les contrôles commenceront pour l'EP. Ce sont les gestionnaires en charge du secteur de l'entreprise profilée (pour en savoir plus sur le calcul de l'APE de l'entreprise à partir des réponses consolidées des UL voir [2]) qui traiteront à partir de cette étape toutes les UL du contour quel que soit leur secteur. Ils réaliseront les contrôles suivants :

- validation de la ventilation en branche du chiffre d'affaires de l'EP ;
- traitement des macro-contrôles de l'EP
- traitement des premiers contrôles (codifications et validation de la ventilation en branche du chiffre d'affaires) des UL de l'EP qui n'ont pas encore été traités et notamment ceux des UL qui répondent après que la phase de macro-contrôle a commencé. Ces contrôles concernent les UL pouvant appartenir à des secteurs différents.

Les micro- et macro-contrôles seront calculés au niveau des UL répondantes et des entreprises ayant dépassé le seuil de couverture du chiffre d'affaires. Ces contrôles se font au niveau du secteur des unités. Ainsi, les micro et macro contrôles des UL des entreprises profilées sont réalisés au niveau de l'APE des UL.

Les micro et macro messages des UL d'un contour ne seront qu'informatifs et sont édités uniquement pour faciliter le travail de contrôle des entreprises profilées par le gestionnaire.

La conséquence de ce mode opératoire de contrôle est que seules les variables de l'entreprise profilée seront contrôlées.

L'exemple ci-dessous est une entreprise constituée de 4 UL (2 UL commerciales, 1 UL industrielle et 1 UL des services) :

	UL 1	UL 2	UL 3	UL 4	EP
Secteur	Commerce	Commerce	Industrie	service	Commerce
répondante	Oui	Non	Oui	Oui	estimé
Ventilation branche	Oui	Oui (extrapolée)	Oui	Oui	somme consolidée des 4 UL
question communes à tous les secteurs	Oui	Oui (extrapolée)	Oui	Oui	somme simple des 4 UL
questions sectorielles commerce	Oui	Oui (extrapolée)	sans objet	sans objet	somme consolidée des UL 1 et 2
questions sectorielles service	sans objet	sans objet	oui	sans objet	sans objet

Les UL 3 et 4 ne recevront qu'un questionnaire « simplifié » car leur partie sectorielle est sans intérêt pour l'entreprise qui appartient au secteur commercial.

À partir de l'exemple ci-dessus, les différents cas de figure suivants peuvent être envisagés :

- si l'UL 2 représente plus de 30 % du chiffre d'affaires, l'entreprise ne sera pas « éligible » pour les macro-contrôles dans un premier temps. Elle ne sera contrôlée qu'après début novembre, lorsque les macro-contrôles seront exécutés sur l'ensemble des entreprises profilées quel que soit le poids des UL non-répondantes ;
- Sinon, si seul un macro contrôle est actif pour l'UL 4, cette dernière ne sera pas à contrôler par les gestionnaires ;
- Si au moins un macro contrôle est actif pour l'EP cette dernière sera à contrôler par un gestionnaire de la division commerce correspondante.

Lorsqu'une entreprise profilée sera à traiter pour un macro-contrôle le gestionnaire pourra :

- corriger directement une des UL du contour s'il constate un problème à ce niveau. La correction faite au niveau de l'UL sera « reportée » au niveau de l'entreprise par le biais du programme de consolidation. Les macro et micro messages au niveau des UL du contour faciliteront la tâche du gestionnaire en mettant en évidence des données a priori atypiques au niveau UL ;
- corriger le flux intra entreprise calculé par le programme de consolidation automatique pour les variables non additives. Cette correction pourra être faite directement en modifiant la valeur du flux. Ces informations sur les flux seront obtenues par le gestionnaire soit après un examen de compte consolidés soit après appel de l'unité légale principale de l'entreprise profilée.

Cette phase de contrôle est suivi d'une phase de traitement de la non-réponse totale avant la diffusion des résultats[3][4].

5. Conclusion

La méthode de consolidation automatique permet donc de calculer des comptes consolidés pour l'ensemble des entreprises non profilées manuellement. La méthode est certes fruste mais elle a l'avantage d'exploiter les données des UL qui sont disponibles sans augmenter la charge statistique des entreprises. L'objectif de notre consolidation est de parvenir à un montant de CA consolidé égal au « vrai » CA. Jusqu'ici on sommait les CA des UL, ce qui nous conduisait à surestimer la valeur réelle du CA de la cible 2. Désormais l'estimation du CA sera plus proche de la réalité, même si cela conduit dans de rares cas à surestimer la consolidation sur certaines entreprises.

Toutefois, nous avons pu constater que nos hypothèses étaient la plupart du temps insuffisantes et nous conduisent à sous-estimer le flux intra groupe. Ce constat est d'autant plus vrai lorsque la

structure du groupe se complexifie. Plusieurs pistes vont être explorées pour améliorer à terme cet algorithme :

- utiliser une enquête spécifique pour demander à l'entreprise ses flux intra (donc interroger la tête de groupe), dans le cas des plus grands groupes non profilés manuellement
- utiliser les enseignements de la première année de campagne pour corriger des binômes d'APE qui n'établiraient aucun flux entre UL en réalité
- bénéficier de l'expertise des gestionnaires à la suite de leurs entretiens avec les groupes
- mettre en place du desk profiling qui consiste à analyser (manuellement dans un premier temps) les rapports d'activité des groupes afin de trouver s'il existe le chiffre d'affaires consolidé France du groupe
- se baser sur l'information la plus fine : les branches d'UL et non de l'information agrégée au niveau de l'activité principale de l'UL: dans un premier temps on envisagera un traitement des branches auxiliaires. Par ailleurs cette modification ne pourra concerner que les entreprises profilées pour lesquelles on disposera d'une réponse à l'ESA. Soit un peu moins du quart des EP et il faudra mesurer l'impact de cette différence de traitement entre les EP échantillonnées et les autres.
- mobiliser l'information sur les ventes intra-groupes issues du cadre clientèle des UL (information qui sera collectée dans la plupart des secteurs de l'ESA) : la question sur les échanges intra-groupes sera présente sur les tous questionnaires de l'ESA à partir de la collecte de l'exercice 2018. La première année sera consacrée :
 - à l'examen de la qualité de réponse des unités légales à cette question
 - à une comparaison entre le chiffre d'affaires consolidé par l'algorithme et celui calculé en utilisant ces nouvelles variables.

Une nouvelle version de l'algorithme prenant en compte cette information pourrait donc être mise en production à partir de l'exercice 2019.

Références

- [1] Le Gleut R., "L'impact du profilage sur la refonte du plan de sondage des Enquêtes Sectorielles Annuelles", *Acte des Journées de Méthodologie Statistique de l'Insee, Paris, 2018*
- [2] Chanteloup G., "The calculation of the automatically profiled enterprises characteristics in the SBR", Unece, Septembre 2017
- [3] Fizzala A., "Comment redresser un échantillon d'unités légales tirées via leurs entreprises ?", *Acte des Journées de Méthodologie Statistique de l'Insee, Paris, 2018*
- [4] Fizzala A., "La gestion par partage des poids des changements de contour des entreprises dans l'Enquête Sectorielle Annuelle", *Acte des Journées de Méthodologie Statistique de l'Insee, Paris, 2018*