The computation of the automatically profiled enterprises characteristics

Workshop statistical unit "enterprise"

Guillaume Chanteloup & Emmanuel Gros

Business statistics directorate

Insee

INSEE
INSTITUT NATIONAL
DE LA STATISTIQUE
ET DES ÉTUDES
ÉCONOMIQUES

Outline of the presentation

1. The context of profiling at Insee

2. Perimeter of the enterprises

3. Computation of the continuity

4. Computation of main activity code

5. Computation of the consolidated turnover

    5.1. Description of the algorithm

    5.2. Some results

6. Computation of other variables

7. Further improvments

Workshop statistical unit "enterprise"

9 & 10 September 2019

The context of profiling in France

➤ French business profiling team: created in 2009, 11 persons today

➤ A profiling strategy depending on size of the groups → 3 different targets:

- Target 1: ~50 of the largest groups (~18% of value added) → progressively profiled by face to face profiling:

  ↳ Enterprises profiled by top-down delineation; tax and survey data tailor-made by profiling team in close co-operation with these groups;

- Target 2: small and medium groups, or simple groups (less than 3 LUs in France) (~100 000 groups) → automatic profiling:

  ↳ Group as a whole = one enterprise;

  ↳ Tax and survey data obtained thanks to consolidation algorithms applied to LUs data.

(~56% of value added)

- Target 3: intermediate-size or -complexity groups (~5 000 groups) → in the short-term, treated like target 2 groups; in the mid-term, semi-automatic profiling based on the results of a survey

Independant legal units account for ~28% of value added

INSEE

Perimeter of the enterprises (1)

➢ The LEUs of the perimeter are split into 2 parts:

- ▪ The LEUs under SBS scope excluding agricultural or financial sector (except 64.20 & 66) and non-market units. This subset forms the ENT perimeter;

- ▪ The other market LEUs if any (belonging to agricultural or financial sector) are each considered as an ENT.

➢ This decision goes against the autonomy criteria put forward elsewhere, so why are we doing this:

- ▪ specificities of financial and insurance activities: specific regulations, different accounting variables, etc.

- ▪ In France, the financial sector is treated by the French national central bank, and the agricultural sector is treated by the ministerial statistical department and out of the SBS scope : elimination of the agricultural/financial units to prevent double counting.

Workshop statistical unit "enterprise"
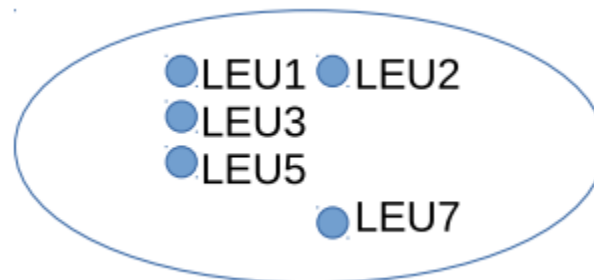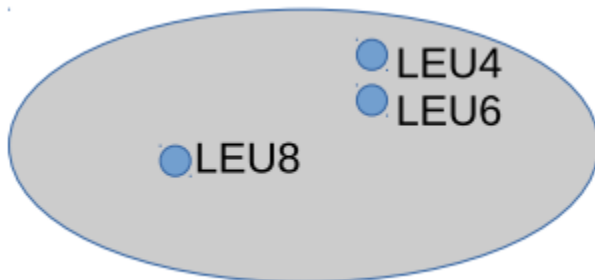
Perimeter of the enterprises (2)

➤ All the LEUs are under SBS scope $\Rightarrow$ 1 enterprise:
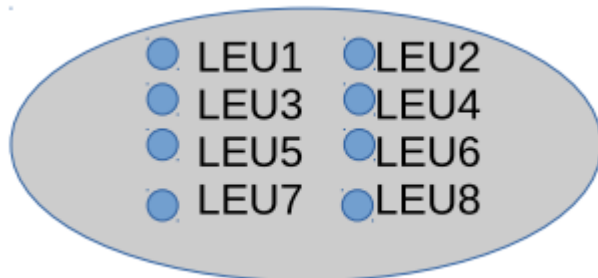
➤ None the LEUs are under SBS scope $\Rightarrow$ LEU=enterprises:
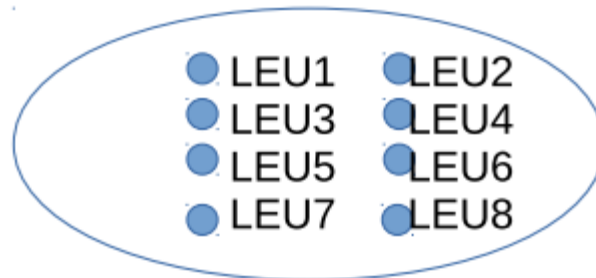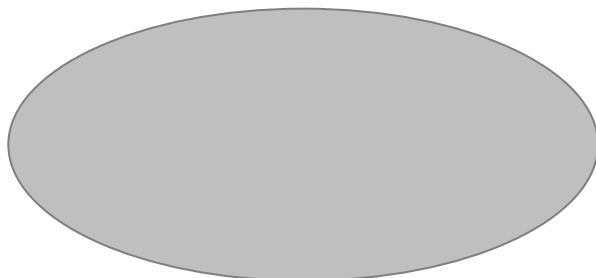
➤ A portion of the LEUs are under SBS scope $\Rightarrow$ 1 enterprise + LEU=enterprises:

Identification and computation of the continuity

➢ The continuity is crucial for the identification of the profiled enterprises and for the dissemination of consistent longitudinal statistics.

➢ For a reference year T, a draft version of the perimeter is given in September T+1, and a final version in April T+2.

➢ For each data supply, a continuity of ENT is calculated:

- ▪ If the ENT is the continuation of the previous one, it keeps its ID;

- ▪ If not, a new ID is provided.

Workshop statistical unit "enterprise"

INSEE

Computation of the continuity

➢ We use the same definition for the continuity of groups and for ENT:

↳ A group (ENT) in T is the continuation of a preliminary group (ENT) in T-1 if and only if the perennial part of these 2 groups (ENTs) is bigger in terms of persons employed than 50 % of the number of persons employed in each one.
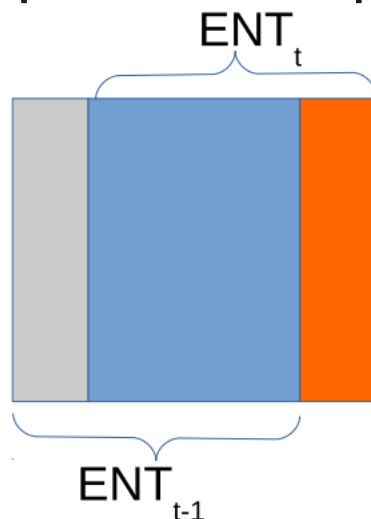
➲ For the year T, employment of $ENT_{t-1} \cap ENT_t > 50$ % of employment of ENT_t…

… and for year T-1, employment of $ENT_t > \cap ENT_{t-1}$ 50 % of employment of ENT_{t-1}

Workshop statistical unit "enterprise"

Computation of the main activity code (1)

➢ Each ENT stored should have a Nace code.

➢ First, they are automatically initialized in the SBR using the turnover and the Nace code of the LEUs of the perimeter of the ENT (next slide).

➢ Then ENTs' Nace code will be updated through their answer to SBS survey (if they are selected in the sample), thanks to the breakdown of the consolidated turnover by activities.

➢ The initialization of ENTs' Nace code should be as close as possible of the Nace code that will be calculated from ENTs' activities breakdown.

Workshop statistical unit "enterprise"

Computation of the main activity code (2)

➢ First step: bottom-up approach establishing the main activity sector of the ENT:
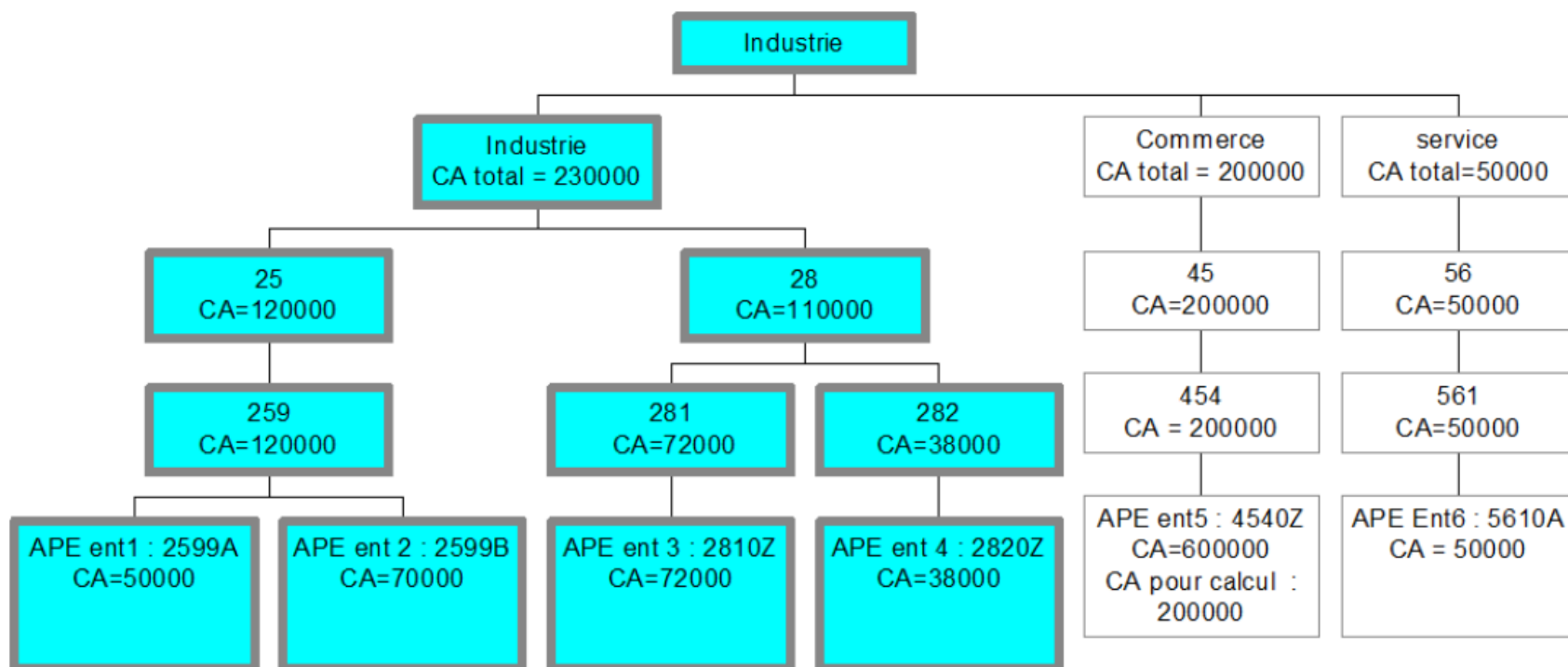
➢ Second step: top-down approach to obtain the Nace code of the ENT:
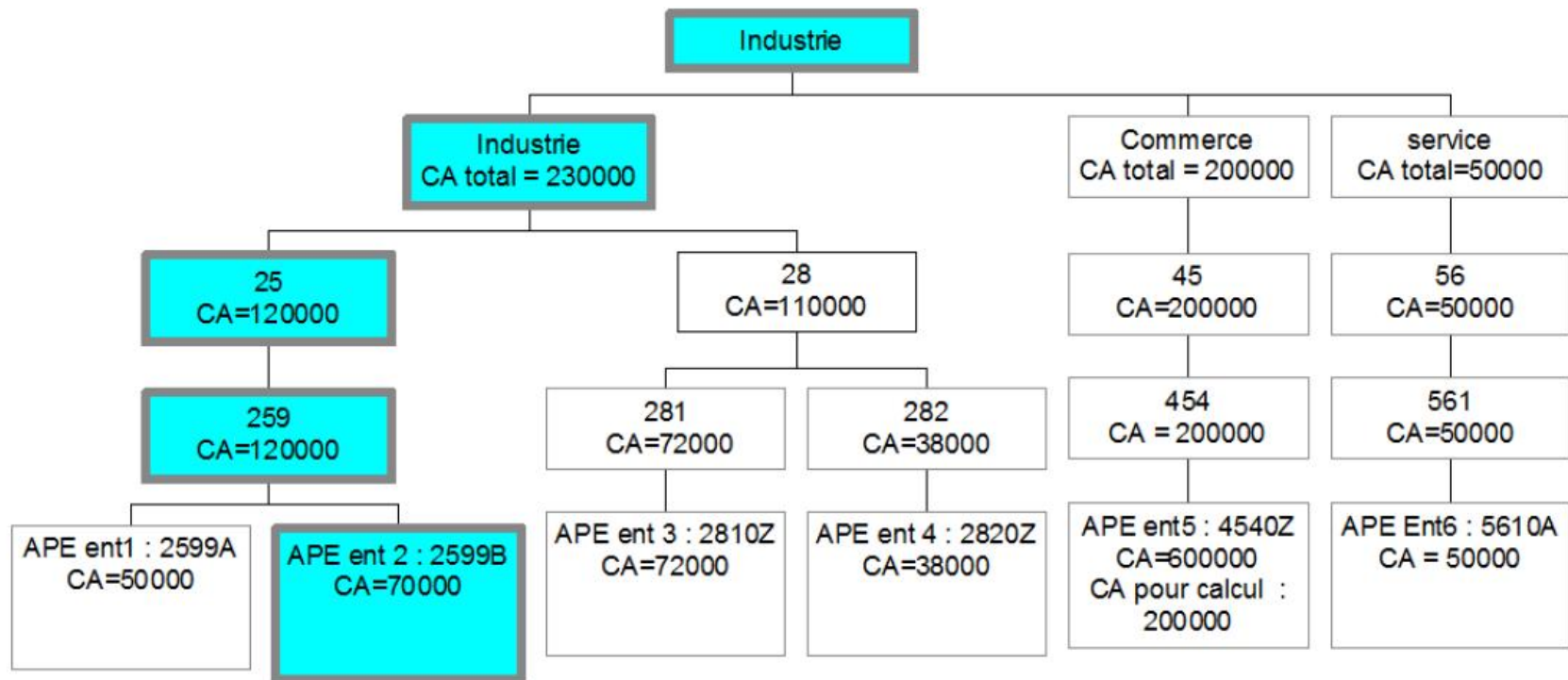
➔ The chosen Nace code is not the activity with the biggest turnover…

Workshop statistical unit "enterprise"

Computation of the consolidated turnover (1/3)

➢ For a given perimeter of an ENT:

❶ Ancillary, commercial or productive type is given to each LEU:

- Commercial: Nace code between 45 and 47;

- Ancillary: Nace code belongs to a list of activities, and the employment in the LEU among the enterprise is lower than an *a priori* threshold – for example sector 64.20 (holding) is always considered as ancillary, whereas 71.10 (architectural and engineering activities) is ancillary only if the employment account for less than 10% of the total ENT employment...

- Productive: if nor commercial, nor ancillary.

Productives

X   X
   X

X   X

Commerciales

X   X   X
   X   X
X

   X   X

Auxiliaires

   X

      X

INSEE
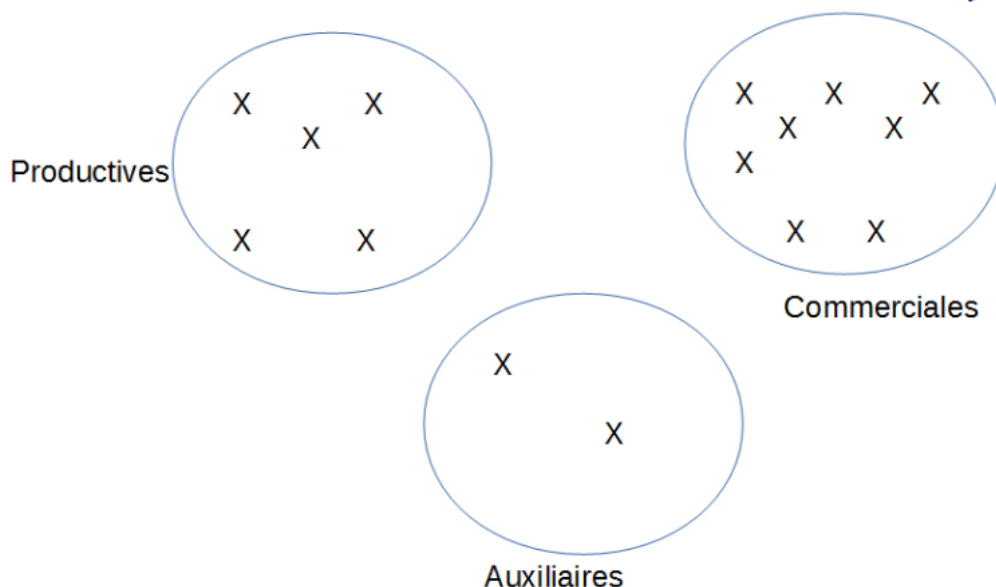
Computation of the consolidated turnover (2/3)

❷ Then, inside each group of productive LEUs and commercial LEUs, we use 2 lists of « *a priori* » integrated activities to establish which LEUs sell to which ones. Thus we can consolidate the sub-groups.
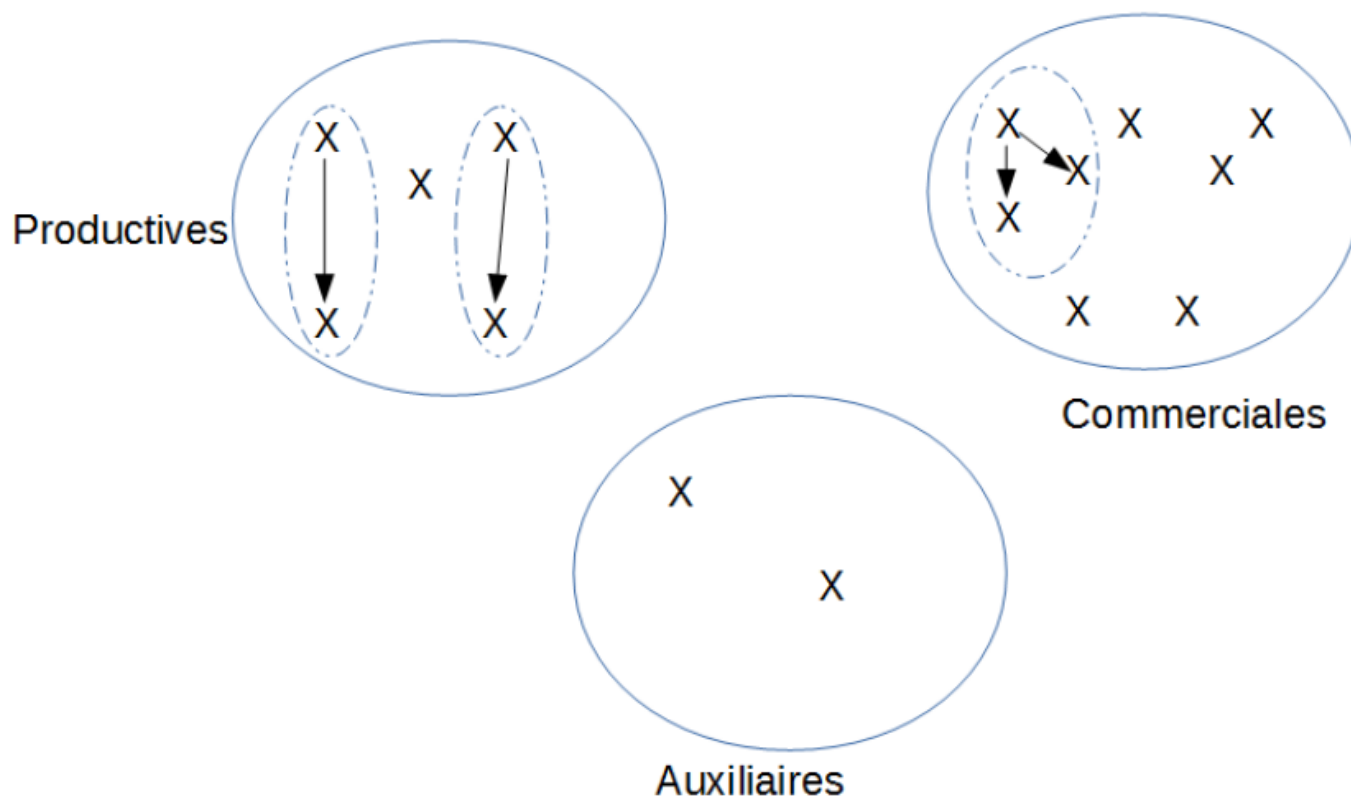
Computation of the consolidated turnover (3/3)

❸ Then, with 2 other lists of *a priori* integrated activities (productive towards commercial and commercial towards productive) to establish which LEUs sell to which ones. Thus we can consolidate the sub-group P+C, and C+P.
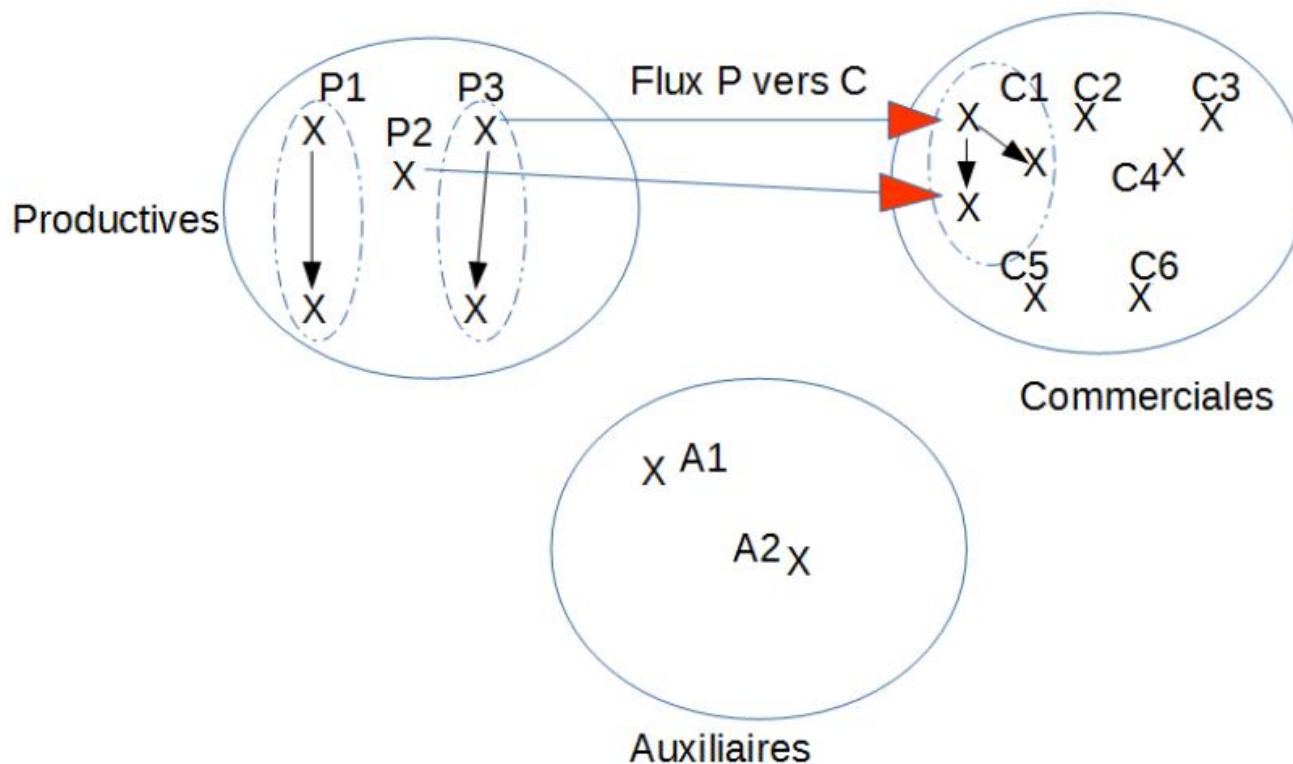
❹ Finally we consolidate all the turnover of the ancillary LEUs

Example of consolidated turnover (1/4)

❶ The following table presents the tax return for 15 LEUS of an ENT, with their typology P, C and A:

	type_UL	Sub-group	Purchase of merchandises	Purchase of raw materials	Sell of goods	Sell of merchandises	Sell of services	Turnover
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

Total	2 301

INSEE

Workshop statistical unit "enterprise"

Example of consolidated turnover (2/4)

❷ Consolidation of the « sub-groups » P1, P3 & C1:

- P1: -130
- P2: -190
- C1 : -170 -50 = -220

	type_UL	Sub-group	Purchase of merchandises	Purchase of raw materials	Sell of goods	Sell of merchandises	Sell of services	Turnover
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

Workshop statistical unit "enterprise"

INSEE

Example of consolidated turnover (3/4)

❸ Consolidation of the « sub-groups » {P2+P3} → C1:

- The sales of goods and merchandises are withdrawn : 17 + 25 + 238 = 280

	type_UL	Sub-group	Purchase of merchandises	Purchase of raw materials	Sell of goods	Sell of merchandises	Sell of services	Turnover
1	P	P1	0	110	*6*	0	0	6
2	P	P1	0	*0*	157	0	0	157
3	P	P2	0	15	**17**	0	0	17
4	P	P3	0	180	**25**	0	0	25
5	P	P3	0	*0*	**238**	0	0	238
6	C	C1	285	0	0	*75*	0	75
7	C	C1	*0*	0	0	198	0	198
8	C	C1	*0*	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

Workshop statistical unit "enterprise"

Example of consolidated turnover (4/4)

❹ Consolidation of the ancillary LEU:

- ▪ Their turnover is subtracted : -100 -200 = -300
- ▪ The intra flow of this ENT is -1,120. The consolidated turnover is 2 301 - 1 120 = 1 181

	type_UL	Sub-group	Purchase of merchandises	Purchase of raw materials	Sell of goods	Sell of merchandises	Sell of services	Turnover
1	P	P1	0	110	*6*	0	0	6
2	P	P1	0	*0*	157	0	0	157
3	P	P2	0	15	*0*	0	0	0
4	P	P3	0	180	*0*	0	0	0
5	P	P3	0	*0*	*0*	0	0	0
6	C	C1	*5*	0	0	*75*	0	75
7	C	C1	*0*	0	0	198	0	198
8	C	C1	*0*	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**	0
15	A	A2	0	0	0	0	**200**	0

	Total	1 181

INSEE

Results of the consolidated turnover

➢ Consolidation effect on all enterprises in billion of €, tax data 2017:

	Intra turnover
A	-43,4
P→C	-64,2
P→P	-6,7
C→P	-1,1
C→C	-7,8
Total	-123,2

➔ Consolidation effect : -5.5 %. Half of the intra flows explained by the productive → commercials ones.

➢ Conservative assumptions – consolidation effect ~12% for manually profiled companies – to avoid over-consolidation...

INSEE

Computation of other variables

➢ For other non additive variables – internal dividends/financial revenues, investments/total equity & trade receivables/trade suppliers –

↳ algorithms based on the subsidiary's shares have been developped;

➢ For additive variables, *e.g.* employment:

↳ sum of the employment of the LEU in the perimeter of the ENT.

Further improvements

➢ Improve the delineation of the automatic profiling: take into account the IPT's results (have more than one ENT in a group).

➢ Improve the quality of the consolidation process:

- A new survey for the largest groups which are not manually profiled;

- Take into accounts new variables of the SBS surveys;

- Use tax data about the enterprise groups;

- Test webscraping.

Workshop statistical unit "enterprise"

Thanks for your attention !

Insee
18 bd Adolphe-Pinard
75675 Paris Cedex 14

www.insee.fr

Informations statistiques :
www.insee.fr / Contacter l'Insee
09 72 72 4000
(coût d'un appel local)
du lundi au vendredi de 9h00 à 17h00

Workshop statistical unit "enterprise"