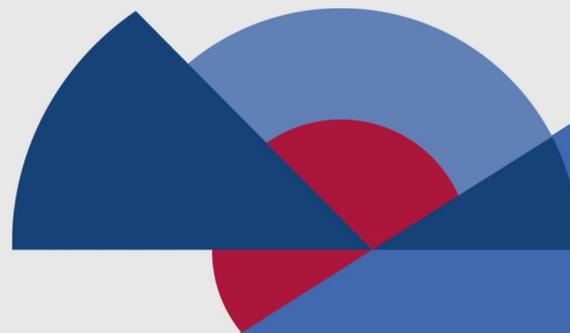


# Passage en entreprises : Algorithme de consolidation et impacts sur le processus Esane

---

Séminaire de Méthodologie Statistique  
20 novembre 2019



Claire JACOD – Division Esane



**01**

• Méthodologie de consolidation

---



**02**

• Impacts et validation de l'algorithme

---



**03**

• Impacts sur le processus Esane

---



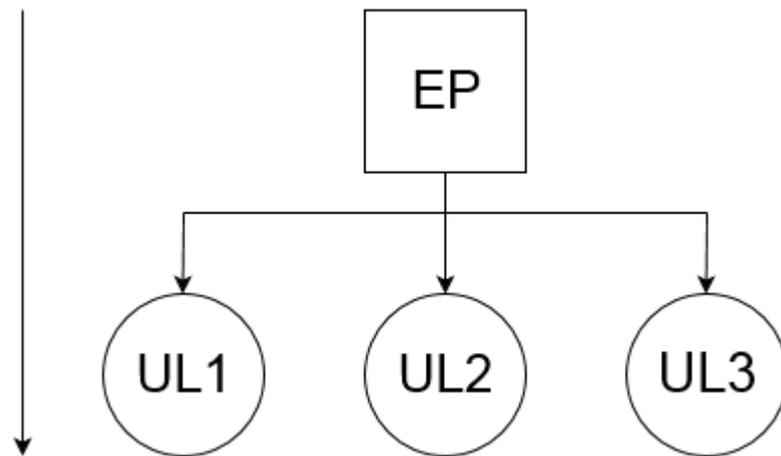
**04**

• Avantages, inconvénients et pistes d'amélioration

---

## Profilage automatique et consolidation

- Une approche top-down pour définir le contour français des entreprises profilées (EP)



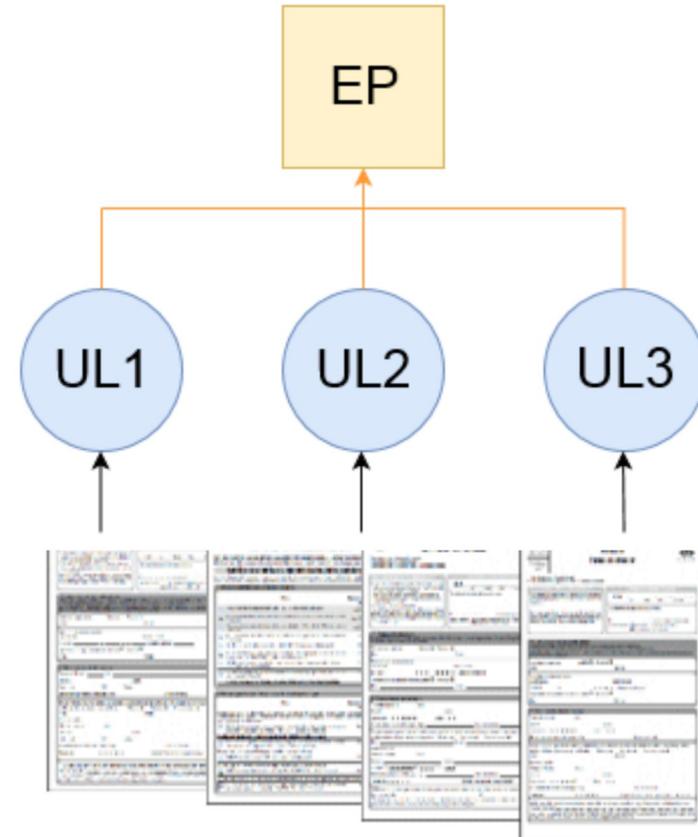
## Profilage automatique et consolidation

- Une approche bottom-up pour construire les comptes des EP

unité statistique

processus de consolidation

unité collectée



# 01

## Méthodologie de consolidation

---



## Principes généraux

On utilise une méthode bottom-up, qui vise à établir le compte de l'EP à partir :

- des réponses des UL du contour (liasses fiscales, enquêtes, données réconciliées)
- des opérations réciproques des UL entre elles : notion de flux

Le chiffre d'affaires consolidé d'une EP s'obtient donc :

- en cumulant les données des UL de son contour
- en corrigeant les postes impactés par ces flux

Mais comment ces flux sont-ils calculés ?

## Détermination de la typologie

Etant donné le contour d'une EP, on définit une typologie pour chaque UL selon son APE :

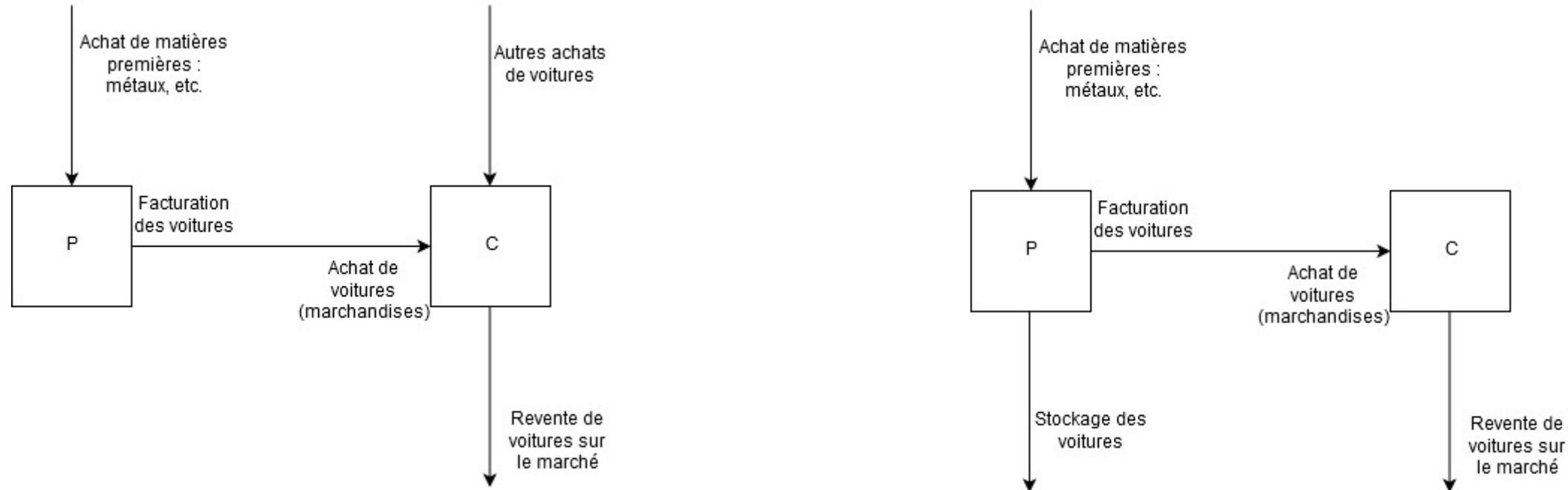
- UL commerciale : activité dans le commerce
- UL auxiliaire :
  - activité appartenant une liste pré-définie (holdings, transport routier de marchandises, etc.)
  - part d'emploi dans l'entreprise inférieure à un seuil pré-défini
- UL productive : les autres UL

## Hypothèses sous-jacentes pour la détermination des flux

- Lorsqu'une UL est auxiliaire, elle n'accède pas au marché et ne sert que l'entreprise
- Utilisation de listes d'activités pour identifier les liens potentiels entre UL : activités intégrées (fabrication de voitures va vendre à commerce de voitures)
- Dans le cas d'activités intégrées :
  - l'UL en amont vend au maximum de ses capacités de vente
  - l'UL en aval achète au maximum de ses capacités d'achats
- Quand une UL n'est ni auxiliaire, ni intégrée, on suppose qu'elle n'échange pas avec les autres UL du contour de son EP

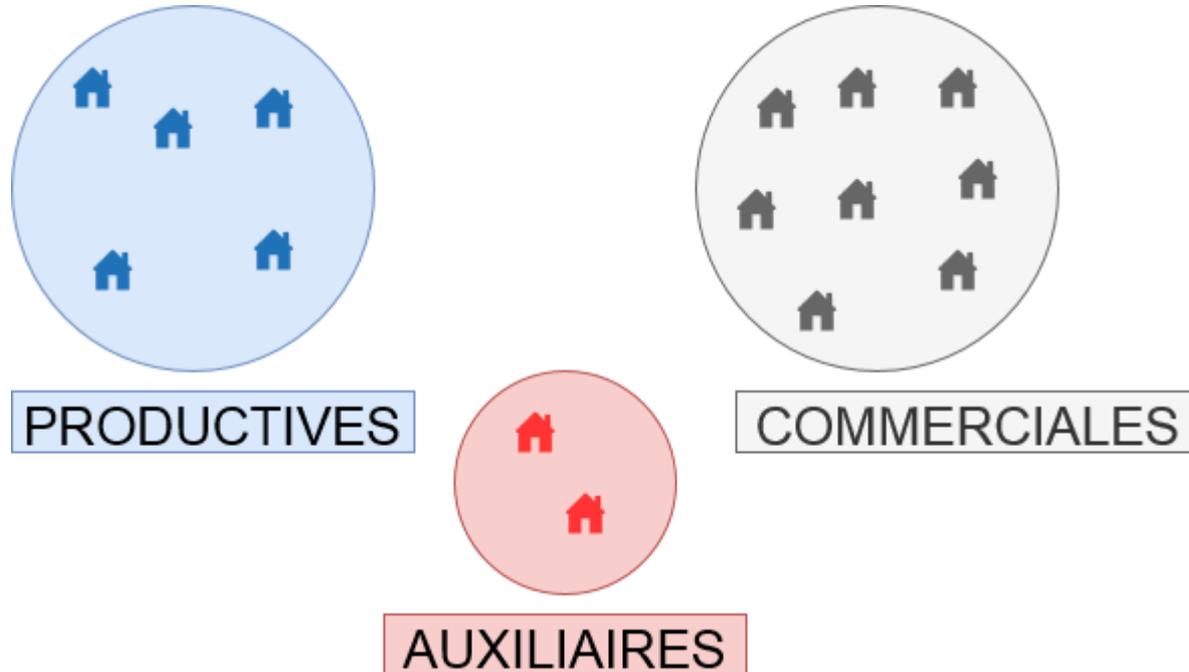
## Calcul des flux

### Principe du calcul des flux entre 2 UL intégrées : un exemple



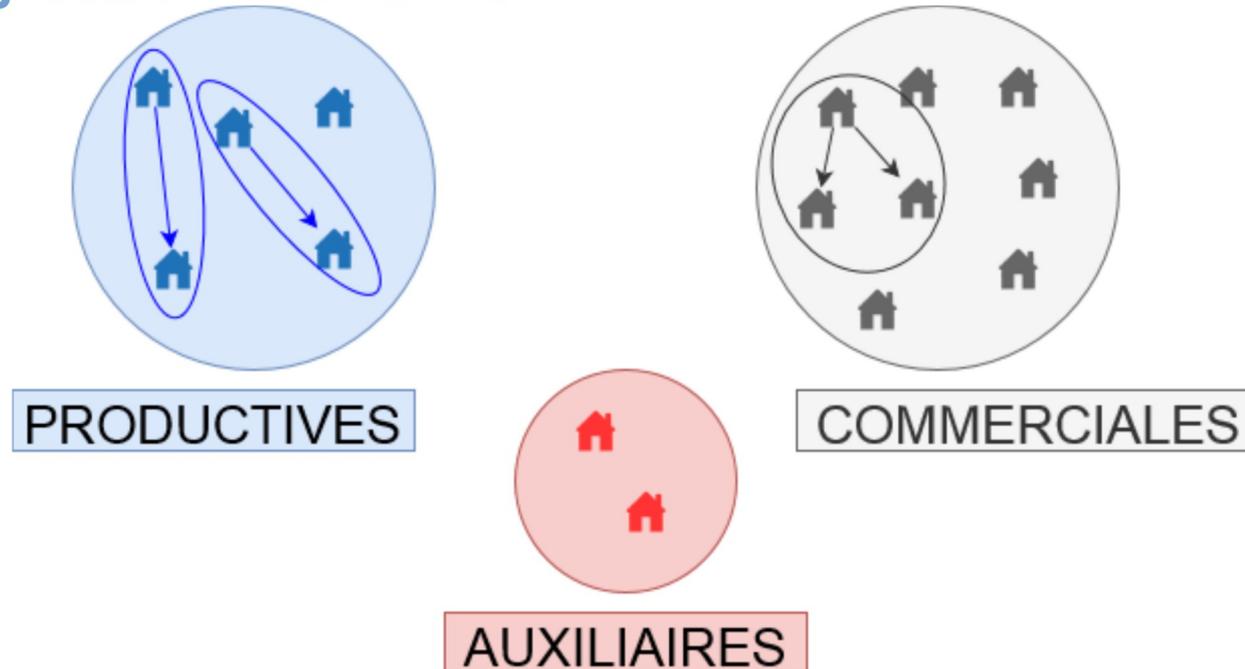
## Calcul des flux

- Etape 1 : déterminer le caractère Productif, Commercial, Auxiliaire (PCA)



## Calcul des flux

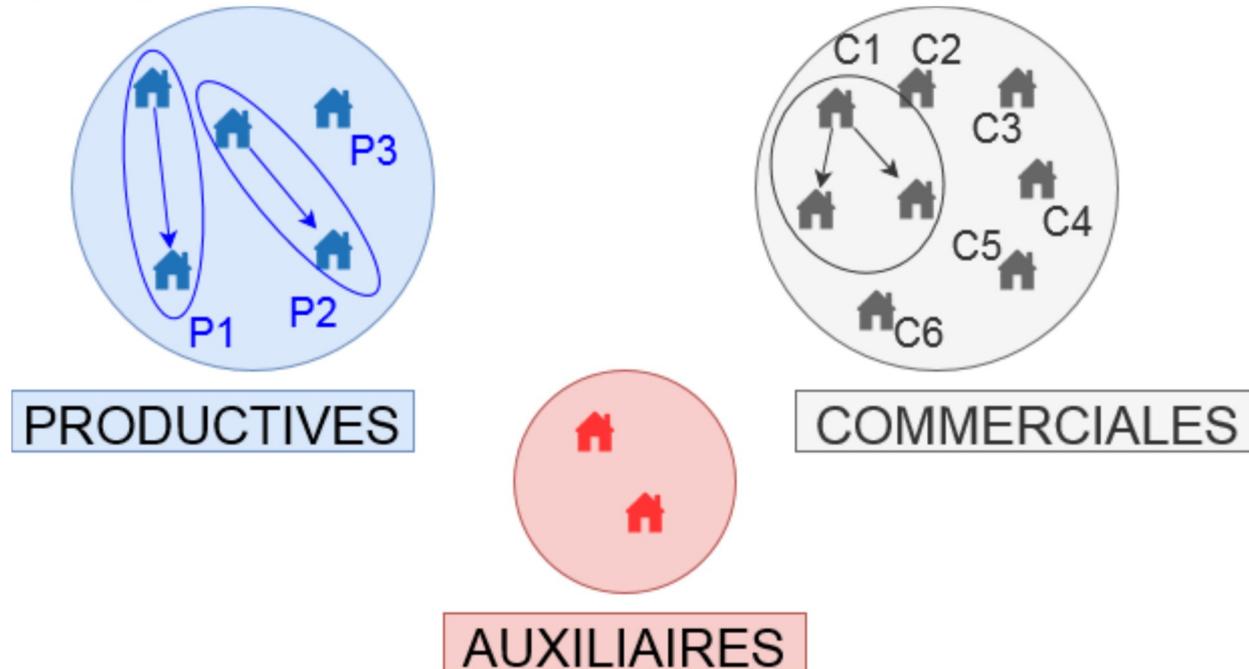
Etape 2 : établir les liens de rang 1 entre UL à partir de la liste des activités intégrées  $P \rightarrow P$  et  $C \rightarrow C$



## Calcul des flux

Etape 2 : établir les liens de rang 1 entre UL à partir de la liste des activités intégrées  $P \rightarrow P$  et  $C \rightarrow C$

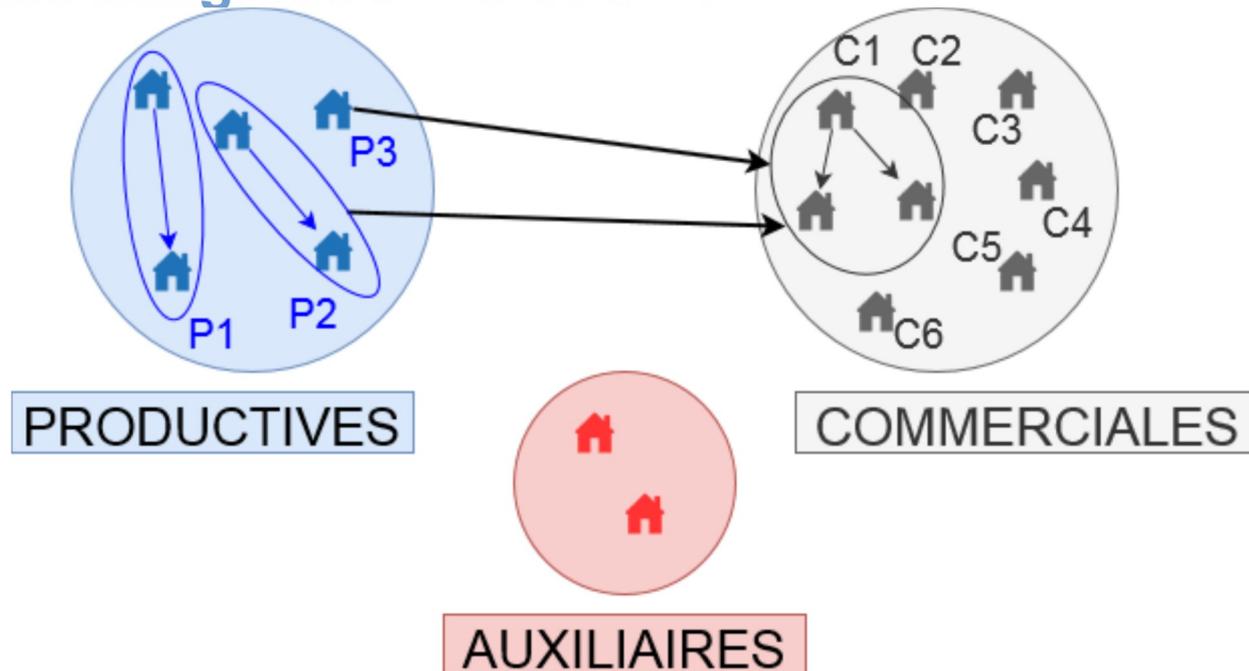
On obtient ainsi des ensembles d'UL en lien (ou sous-EP), d'autres qui ne le sont pas



## Calcul des flux

Etape 3 : établir les liens de rang 2 entre UL et/ou des sous-EP à partir de la liste des activités intégrées  $P \rightarrow C$  et  $C \rightarrow P$

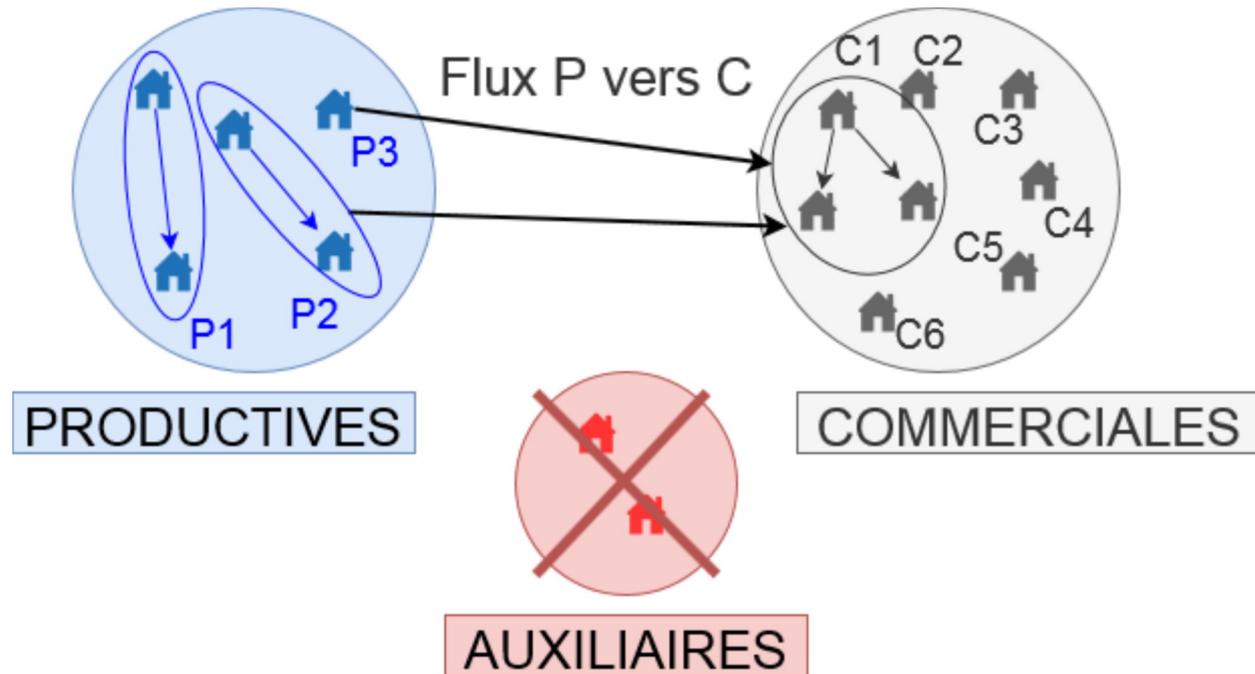
Ici, les ensembles P2 et P3 seront consolidés avec C1



## Calcul des flux

### Etape 4 : consolider le chiffre d'affaires des UL auxiliaires

Suppression de tout ou partie du CA des UL auxiliaires



# 02

## Impacts et validation de l'algorithme

---



### Impact de la consolidation calculée par l'algorithme

La moitié du flux total est expliqué par les flux entre productives et commerciales

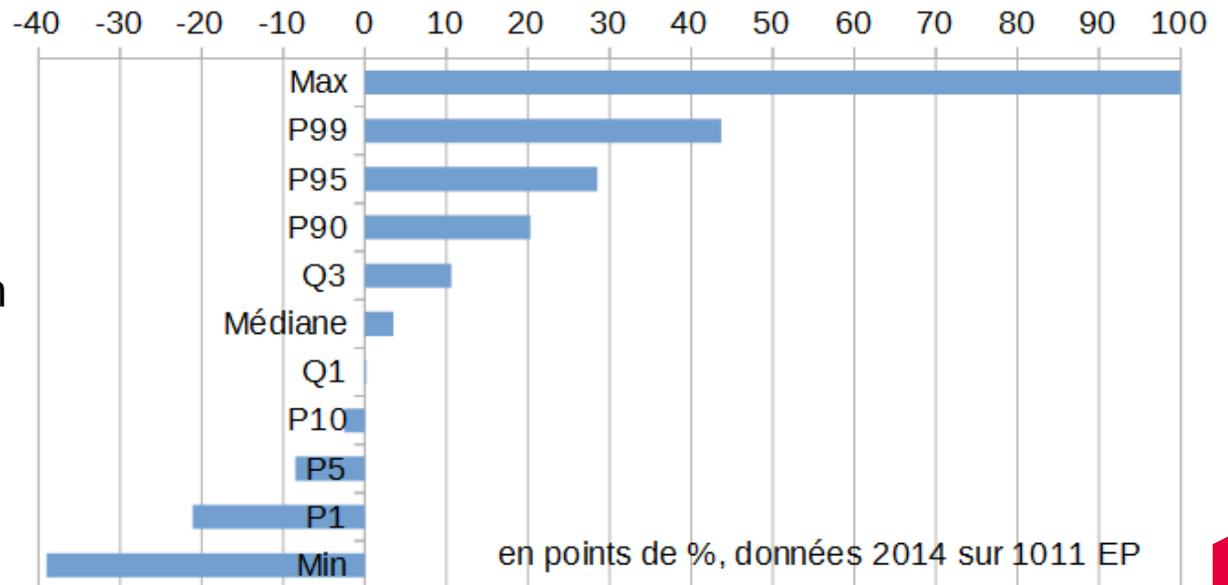
Typologie	2017 – Flux de CA
A	35%
P → C	52%
C → P	4%
P → P	4%
C → C	5%
Total	121 Mds €

### Validation de l'algorithme

#### Comparaison des résultats de l'algorithme avec les données de la Banque de France

- Pour 80% des EP, le CA consolidé par l'algorithme est proche du "vrai" CA ou supérieur
- Dans la plupart des cas, on sous-consolide : préférence initiale pour la sous-consolidation respectée
- Cas de surconsolidation : pistes d'amélioration sollicitant les données de l'ESA

Ecarts entre le taux de consolidation Banque de France et le taux de consolidation calculé par algorithme



# 03

## Impacts sur le processus Esane

---



### Evolution majeure de l'ensemble du processus Esane

- Pour le plan de sondage des enquêtes ESA et EAP
- **Pour le traitement et le contrôle des données (enquête, données fiscales, réconciliation des données) :**
  - **Pré-requis aux contrôles : contours et initialisation N-1**
  - **Traitements automatiques et manuels**
- Pour la diffusion

## Evolution majeure de l'ensemble du processus Esane

### Principes de contrôle des données

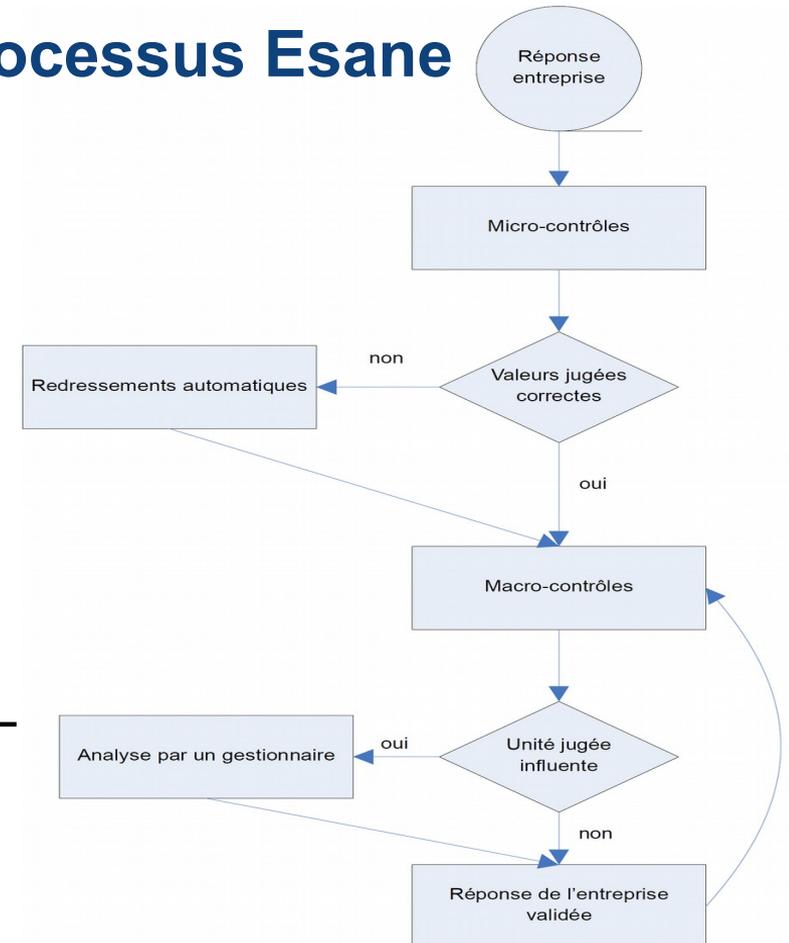
AVANT :

Contrôles

Qualification

Redressements

Influence des UL calculée sur des agrégats en UL



# Evolution majeure de l'ensemble du processus Esane

## Principes de contrôle des données

**MAINTENANT** : même logique, mais pilotage des contrôles en entreprises

Influence des unités calculée sur des agrégats en entreprises

Contrôles des unités :

- aucun changement pour les UL indépendantes
- pour les entreprises profilées automatiquement : modification des traitements automatiques et des écrans et consignes de gestion

## Evolution majeure de l'ensemble du processus Esane

### Principes de contrôle des données

MAINTENANT, pour les EP profilées automatiquement :

1. Contrôle, qualification et redressement des UL de contour

2. Consolidation : données consolidées =  $\sum_{i \in EP} \text{données de l'UL}_i$  - « flux intra »

*(A blue bracket underlines the summation term, with the text "nécessité d'un contour complet => imputation des unités manquantes" below it.)*

3. Contrôle, qualification et redressement des EP :

- seuil de lancement des contrôles en EP : critère de réponse/qualité sur les UL de contour
- les EP influentes sortent à traiter : le problème peut alors venir :
  - des données des UL
  - et/ou
  - des flux intra estimés, avec un contrôle spécifique des flux importants (pour éviter la surconsolidation)

### Evolution majeure de l'ensemble du processus Esane

#### La gestion

Logique de traitement d'une EP :

- **contrôler et corriger si besoin les UL** de l'EP qui ressortent en contrôle manuel d'après les résultats des macro-contrôles en UL
- si une de ces UL est **corrigée**, **vérifier l'impact** de cette correction **au niveau de l'EP**
- si après traitement de toutes les UL à contrôler, l'EP reste toujours problématique, alors le problème potentiel vient *a priori* du flux estimé  
=> **valider/corriger le flux**

# Evolution majeure de l'ensemble du processus Esane

## La gestion

Conséquences sur les UL de contour d'entreprises profilées automatiquement :

- seules celles appartenant à une EP à contrôler pourront être contrôlées manuellement
- les UL influentes appartenant à des EP non influentes ne seront plus contrôlées manuellement

Quelques contrôles sont (aussi) réalisés en UL : gestion de la collecte ESA, contrôles des liasses en doublons, validation de l'APE...

# Evolution majeure de l'ensemble du processus Esane

## Contours des EP et initialisation N-1

Une diffusion et une gestion en couple N/N-1 : les groupes et donc les EP **changent de contour** :

> **QUESTION** : comment traiter manuellement des unités influentes sur les évolutions N/N-1 ?

- Problème de contour ?
- Problème d'UL ?
- Problème de flux ?

=> Pour les entreprises profilées automatiquement, **on travaille à contours constants** : contours N pour N-1, pour des contrôles temporels plus pertinents

Pour les entreprises profilées manuellement, méthode inchangée / avant :

- difficulté d'obtenir les flux N-1 à contours N => on travaille à contours N pour N et N-1 pour N-1
- sauf changement de méthode de profilage, auquel cas un pro-format N-1 à contours et méthode N est utilisé

# Evolution majeure de l'ensemble du processus Esane

## Contours des EP et initialisation N-1

Conséquences :

- Diffusion :
  - Résultats N-1 à méthodes N et N-1 non comparables
  - Différence de traitement entre les entreprises profilées manuellement et automatiquement
- Production :
  - Chargement des contours et échantillons N et N-1 avant le démarrage de la gestion
  - Initialisation N-1 : recalcul des données N-1 nécessaire dans tous les sous-systèmes avant le démarrage des macro-contrôles, sur la base des contours N

# 04 Avantages, inconvénients et pistes d'amélioration



### Avantages

- **Objectif de diffusion des données 2017 atteint**
  - Pas d'augmentation de la charge statistique des UL
  - Permet d'être plus proche du véritable CA que de sommer simplement les résultats des UL de contour

### Inconvénients

- Algorithme :
  - Sous-estimation globale des flux et surestimation pour certaines EP
  - Fragilité pour les plus grosses EP profilées automatiquement
- Processus Esane : augmentation de la complexité du processus, avec l'ajout de nombreux traitements et de nouveaux concepts

### Pistes d'amélioration

#### - Algorithme

- Correction/amélioration des binômes d'APE
- Utilisation de l'information plus fine sur les branches et non pas l'APE des UL pour la typologie et liens entre UL

### Pistes d'amélioration

#### - Processus

- Pour les plus grosses EP profilées automatiquement :
  - “Light profiling” : collecte d’informations au niveau UL, complétée par des informations de flux collectées auprès du groupe par le profileur => 1<sup>ère</sup> expérimentation sur 2018
  - A moyen terme (2025 ?) : utiliser une enquête spécifique sur les flux intra, étude de faisabilité par le REE en 2020
- Informations complémentaires issues de l’ESA :
  - validité 2018 : part des achats/CA dans le groupe
  - prise en compte des variables d’export
- Priorité : stabilisation du processus actuel de production

## Retrouvez-nous sur :

[insee.fr](https://www.insee.fr)



---

Guillaume Chanteloup  
Ex-CPS Profilage cible 2

Claire Jacod  
Responsable de la section Production et  
maintenance Esane

[Cliquez pour ajouter un titre](#)