

# Formation aux statistiques structurelles d'entreprises

Formation Insee / CASD - Chercheurs



11/03/2022

- 1. Présentation générale de la statistique structurelle d'entreprise**
- 2. Les traitements réalisés sur les données**
- 3. Le profilage des groupes en entreprises**
- 4. Usages et limites des données et de leur utilisation**
- 5. Les fichiers de diffusion et la documentation**
- 6. Exemples d'utilisation des données dans des études à l'Insee**



**01**

# Présentation générale de la statistique structurelle d'entreprise

---



Un **ensemble de dispositifs** qui permet de produire chaque année de nombreuses données sur l'activité des entreprises en France, en réponse à différents besoins

Un **répertoire statistique** (Sirus) pour déterminer la liste des unités, le champ de la statistique d'entreprise, regrouper des informations (Sirene, emploi, variables d'enquêtes, etc.). Utilisé pour les besoins internes (tirage d'échantillon, référentiel des processus de production notamment) => *pas disponible au CASD*

Des outils spécifiques pour connaître les **restructurations** entre unités légales (Citrus) et le **contour des groupes de sociétés (Lifi)** => *fichiers disponibles au CASD*



Un processus spécifique sur la **démographie d'entreprise** (stocks, création, transferts) => *fichiers disponibles au CASD*

Un dispositif pour collecter les **données comptables et financières** des entreprises et **sur leur activité** (dispositif Esane : FARE et autres fichiers) => *fichiers disponibles au CASD*

Des **enquêtes thématiques** : TIC, Siné, Ofats, EACEI, CAM, aéronautique...  
=> *fichiers disponibles au CASD*

**Des limites** : sources administratives pas toujours adaptées aux besoins statistiques, données d'enquêtes limitées par la taille des échantillons, certaines données difficiles à obtenir des entreprises



Système calé sur des **concepts et définitions internationales**, en cohérence avec les données macroéconomiques : champ, unités d'observation, APE...

Données collectées venant de la **comptabilité d'entreprise**

**Système d'information très riche** grâce aux sources utilisées et aux moyens mis en œuvre à l'Insee pour contrôler et produire des indicateurs statistiques

Système conçu pour produire **chaque année la meilleure image** possible du tissu productif (en cohérence avec l'année précédente)

**Refonte en 2008-2009** (projet Résane) : amélioration du processus, données unifiées, production de données sur les « entreprises » au sens de la LME plus récemment

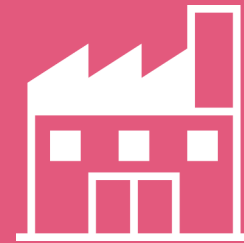
Des **adaptations nécessaires** face aux enjeux actuels



**02**

## Les traitements sur les données

---



## ● ELABORATION DES STATISTIQUES ANNUELLES D'ENTREPRISES (ESANE)

- Utilisation de deux sources principales :
  - Données administratives : liasses fiscales (convention Dgfip, notion de secret)
  - Deux enquêtes auprès des entreprises : ESA et EAP
- Utilisation d'autres sources : Sirius, DSN, Acoess...
- Objectifs :
  - Réduire la charge d'enquête des entreprises
  - Les enquêtes sont donc complémentaires des liasses, avec comme apports principaux :
    - La ventilation par branche de l'activité (connaissance de la totalité des activités et détermination de l'APE)
    - Questions sectorielles : informations non disponible dans les liasses, variables selon l'unité interrogée
  - Principe de réconciliation des données : une seule donnée en sortie, auparavant, plusieurs fichiers en sortie...
- En sortie du dispositif : le fichier FARE...mais pas uniquement (voir point 5)





## ● LES TRAITEMENTS :

### – Un processus assez dense : quelques notes et documents de référence :

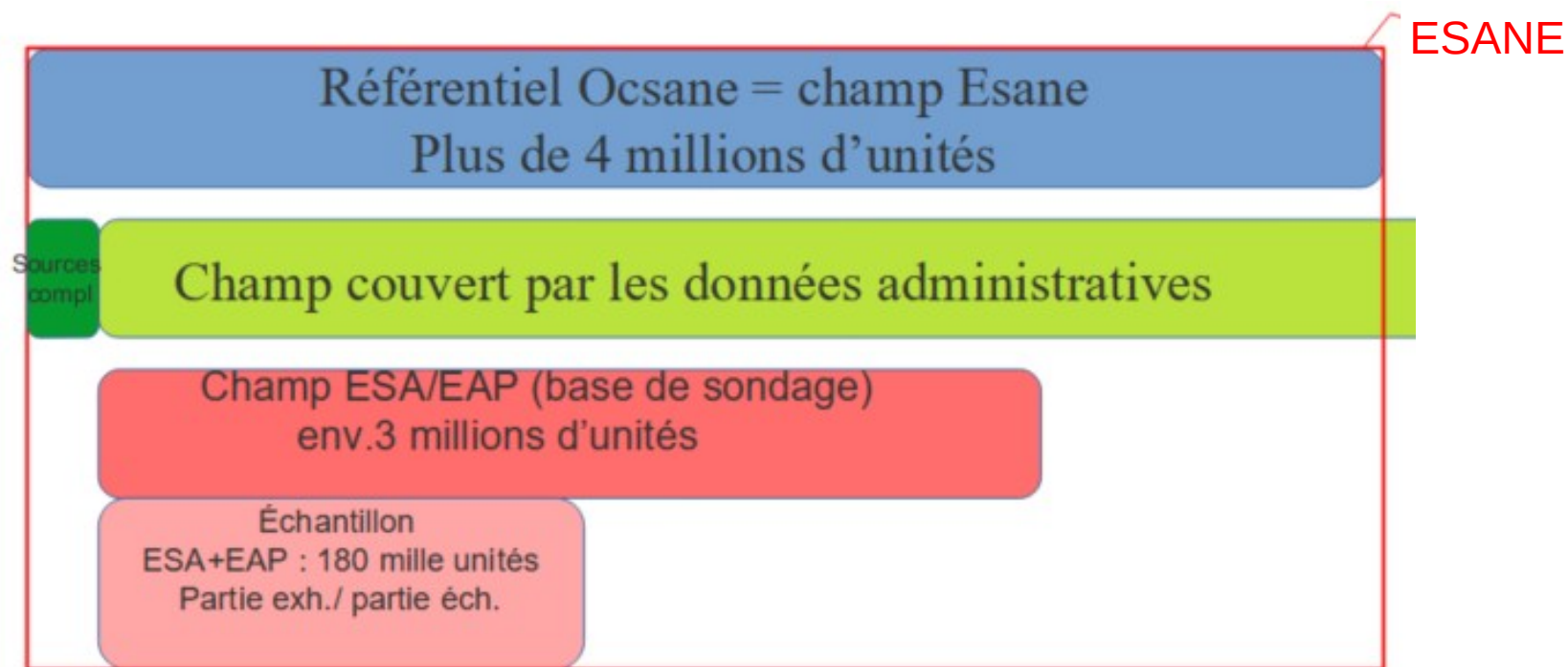
- « Esane méthodes » qui décrit la méthodologie du système Esane
- Les différentes notes de mise à disposition qui accompagne les fichiers tels que le Fare

### – Quelques rappels sur le champ Esane :

#### • Champ :

- sociétés et entreprises individuelles, non agricoles et non financières (exceptés les auxiliaires financiers et d'assurances ainsi que les holdings qui font partie du champ depuis 2010).
- champ géographique : France entière
- comprend les sociétés juridiques considérées dans le répertoire statistique comme marchandes et exploitantes ou participant au système productif.
- Attention, c'est un champ « a priori », élaboré à partir du répertoire Sirius : ce ne sont pas les données administratives reçues qui déterminent le champ et le champ évolue.
- Quelques évolutions récentes : intégration de Mayotte et amélioration de la prise en compte des micro-entrepreneurs au sens fiscal (voir page 79-80 de l'Esane Méthodes pour l'historique)

- Le champ Esane comprend des sous-champs emboîtés : attendu, couvert par les données admin., couvert par les enquêtes etc.



## ● LES TRAITEMENTS :

### – Retour sur la réconciliation des données

- Depuis la refonte, Esane produit un seul jeu de données là où on avait auparavant des systèmes indépendants (EAE, SUSE...)
- Le principe :
  - pour certaines unités et pour certaines variables communes , 2 sources disponibles donc choix de la donnée pertinente pour avoir une seule donnée en sortie.
  - Par ex : on privilégie plutôt les enquêtes pour la ventilation du CA et les liasses pour le CA.

### – Le sujet des imputations :

- Certaines liasses sont imputées (variable présente dans le fichier) car attendues dans le champ
- Le taux global est élevé (35%) mais ne représente que 3 % environ en CA
- ...car se sont surtout des micro-entreprises (imputées par défaut, plus d'1 million d'unités), sinon taux de 13 %



## ● LES TRAITEMENTS :

- **Esane : un dispositif bien contrôlé et redressé en amont de la diffusion :**
  - Production / automatique : sur liasses et enquête, micro et macro contrôles
  - Production / manuel : intervention des gestionnaires (env. 60), hiérarchisé et uniquement sur les unités les plus influentes
  - Après la collecte mais avant la diffusion des données : étapes de « pré-diffusion »
- **Des données N-1 à méthode N pour faciliter les comparaisons :**
  - Fabrication d'un fichier N-1 à champ/méthodologie/contours comparables à N
  - Permet de garantir des contours homogènes en entreprise sur les deux années
  - Permet de corriger/améliorer les données N-1 même pendant la campagne N

# 03

## Le profilage des groupes en entreprises

---



Organisation en groupe de sociétés : **un phénomène croissant** dans un contexte de mondialisation des économies modernes

Unité légale juridiquement indépendante mais **sans autonomie de décision sur le plan opérationnel** au sein d'un groupe

**Echanges commerciaux entre sociétés** d'un groupe ne respectant pas forcément les règles du marché.  
Mouvement des moyens de production et des capitaux par optimisation économique ou financière.

Mais pas d'existence juridique et donc **difficile à prendre compte dans nos statistiques**

Un premier règlement européen en 1993 sur le sujet

Figure 1. La description simplifiée des flux d'une UL (unité légale)

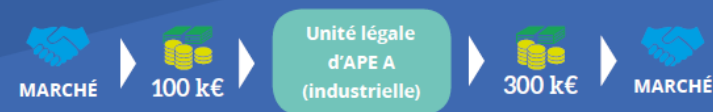
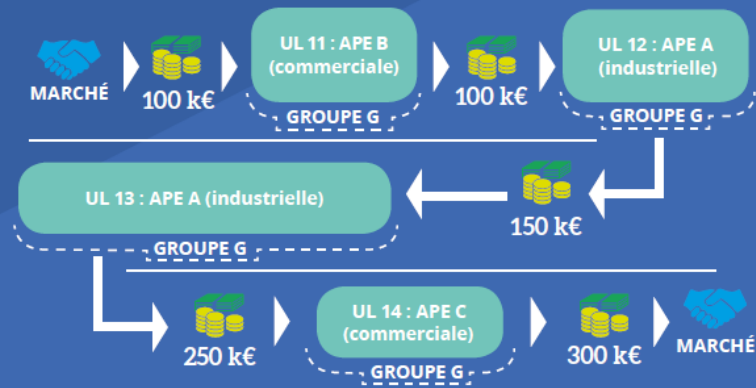


Figure 2. La description simplifiée des flux au sein d'un groupe



Un « **choc** » sur nos données à la fin des années 90 suite à la restructuration de deux grands constructeurs automobiles : un chiffre d'affaires multiplié « fictivement » par trois !

L'« **entreprise** » au sens de la **LME**, unité statistique pertinente pour l'analyse économique (rapport du Cnis, 2008)

Décret 2008-1354 (LME) : l'entreprise est *la plus petite combinaison d'unités légales qui constitue une unité organisationnelle de production de biens et services jouissant d'une **certaine autonomie de décision**.*

**Adaptation progressive de notre système de production** : quelques groupes à partir de la fin des années 90, constitution d'une équipe de profileurs en 2011, développement d'une méthode pour tous les groupes en 2016

Une **vision nouvelle du tissu productif** (taille des unités, réallocations sectorielles, performances des entreprises)



- Déterminer le **contour des entreprises** au sein des groupes :
  - à partir de Lifi sur le champ Esane (yc trace française), analyse de la continuité économique
  - découpage en entreprises par les profileurs pour 60 des plus grands groupes selon le critère d'autonomie de décision
- Identifier les variables non additives (achats, ventes, créances/dettes, emprunts/prêts, provisions, dividendes, titres de participations internes...) et **consolider leurs valeurs** (supprimer le flux intra-groupes et les activités auxiliaires)
- **Produire et contrôler** les données au niveau de l'entreprise (à partir de données collectées essentiellement en UL)



- **Une méthode manuelle pour 60 des grands groupes**, en collaboration avec eux : actualisation du contour des groupes en UL, découpage en entreprise (grille d'analyse pour déterminer l'autonomie de décision), analyse des comptes consolidés du groupe et des données des liasses fiscales des UL (et quelques ESA). Détermination de la ventilation du chiffre d'affaires par branche.
- **Une consolidation automatique pour les autres groupes** : pas de découpage en entreprise. Pour le chiffre d'affaires, identification des flux intra-groupes à partir d'une table de binôme basée sur les codes d'activités des UL, puis consolidation au maximum des ventes et achats selon différentes typologies. Contrôles et améliorations des données influentes.
- **Des pistes d'amélioration** : un travail manuel sur plus de grands groupes multinationaux, un nouvel algorithme de consolidation automatique à partir de nouvelles réponses à l'ESA et l'EAP, l'utilisation de nouvelles sources sur les variables du bilan

# 04

## Usages et limites des données et de leur utilisation

---



## ● USAGES (1/2)

### – Identifiant : Siren

- Pour les UL et aussi pour les entreprises profilées (EP) : commence par P
- Siren EP est disponible pour chaque UL appartenant à son contour
- Appariement avec d'autres sources ayant même identifiant UL (ex. DSN – emploi)
- Suivre les unités dans le temps

### – Permet d'avoir la meilleure information (photo) l'année N

### – Analyser les évolutions

- Utilisation du fichier N-1 à contours et méthode N
- Attention, pour les évolutions des EP, les contours des EP évoluent : si changement trop important, numéro d'identification change
  - **Pas possibilité de retrouver le même siren entre N et N-1 et antérieurement**

## ● USAGES (2/2)

- Étudier des grandeurs économiques
  - par type d'unités (via la catégorie d'entreprise), par exemple
- Étude sur des sous-populations
  - Par exemple, sur les PME, les entreprises de + 20 salariés
- Analyser les activités secondaires au sein des secteurs
- Exploiter les données de l'enquête sectorielle annuelle (ESA)



## ● LIMITES (1/5)

- Identifiant : Siren - Pas de données au niveau établissement (siret)
- Étude régionalisée à faire avec précaution (ou infra-régional) :
  - Plan de sondage non représentatif au niveau régional
  - 100 % activité affectée au siren ( → région du siège social, or les établissements peuvent être répartis dans plusieurs régions)
  - Limite renforcée au passage en entreprise

=> Études possibles sur UL mono-régionales, estimation d'une richesse dégagée en proratisant le CA/VA par l'emploi des établissements, étude de la dépendance des UL aux centres de décision des groupes, etc.

## ● LIMITES (2/5)

- Difficulté de faire des évolutions sur longue période au niveau Entreprise
- Contour des entreprises EP Esane incomplet :
  - restreint aux unités du champ Esane donc hors agriculture (sauf 02) et hors financier (sauf 64.20Z et 66)
- => Intérêt de distinguer dans une étude les UL d'une entreprise des autres UL (comportement différent, déclaration volatile, etc.)
- Depuis millésime 2017, restructurations pas aussi bien prises en compte qu'auparavant
  - Conséquence : niveau de l'investissement est sur-estimé

## ● LIMITES (3/5)

– Depuis millésime 2017, avec le passage progressif à la DSN, l'emploi moins bien mesuré

- Rupture de série et de concept
- Moins d'information disponible

=> Appariement avec DSN préférable

– Données pas complètement comparables avec celles de la comptabilité nationale

- Pas même niveau (passage au concept comptabilité nationale)
- Données comptables assises sur le PCG
- Mais évolutions proches sur les entreprises non financières (CA, VA)

## ● LIMITES (4/5)

- **Déconseillé d'utiliser les données Groupes du FARE pour des études sur les groupes :**
  - Données Groupes proviennent de LIFI N : existence de décalage entre contours des entreprises et les données Groupe
  - Données Groupes sont sur le contour complet (y c. agriculture et financier) – peut générer des incohérences
  - Pas de consolidation faite au niveau groupe (mais au niveau EP)
- => **Mais utilisation possible des caractéristiques des groupes**
- **Information disponible que sur la trace française des groupes**





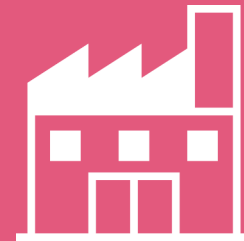
## ● LIMITES (5/5)

- **Information disponible que sur la trace française des groupes**
  - Échanges commerciaux avec sociétés du groupe à l'étranger pas consolidées
  - Déclarations pouvant passer d'un pays à l'autre par optimisation (dettes, brevets, bénéfices...)
  - Entreprises multinationales difficiles à suivre

# 05

## Les fichiers de diffusion et la documentation

---



## ● 3 PRINCIPAUX FICHIERS INDIVIDUELS ESANE (1/3)

- **FARE : Fichier Approché des Résultats Esane (ex FICUS) – à partir de 2008**
  - Données comptables (bilan, compte de résultat)
  - Données des liasses ou issues du processus de réconciliation des données (variables suffixées par redi)
  - Exhaustivité (y c. imputation) et donc sans poids
  - Niveau de diffusion : UL et EP : utilisation des indicatrices diff\_ul ou diff\_ep
  - Pour un millésime, 2 années : N-1 et N à contour et méthode N
  - Disponible en sept.-octobre N+2

**Attention, somme des données individuelles est différente des agrégats publiés sur Insee.fr (estimateur par différence)**

## ● EN COMPLÉMENT DU FARE, AVANT 2008

### – Le fichier FICUS : fichier unifiée de statistiques d'entreprises

- Données disponibles depuis 1994
- Élaboré à partir des mêmes sources que le FARE (en particulier les liasses fiscales)
- Traitements un peu différents (ex : pas de réconciliation du chiffre d'affaires - égalité – mais valeur approchée)
- Champ différent :
  - ne contient que les entreprises imposées au BIC et BNC
  - Mais hors micro-entreprise
  - Exclusion des exploitations agricoles du régime BIC et celles au régime Bénéfices agricoles
- Evolution possible du questionnement et définition des variables (dépend des évolutions fiscales notamment)

## ● 3 PRINCIPAUX FICHIERS INDIVIDUELS ESANE (2/3)

### – VAC : Ventilation du CA par activité

- CA ventilé par branche (5 positions)
- Croisement branche (variable co) et secteur (variable ape\_diff)
- Unités échantillonnées à l'exhaustif ou répondantes à l'ESA et à l'EAP (industrielles)
- Niveau de diffusion : UL et EP : utilisation des indicatrices diff\_ul ou diff\_ep
- 2 jeux de pondération : poids\_apres\_calage\_ul et poids\_apres\_calage\_ep
- Pour un millésime, 2 années : N-1 et N à contour et méthode N
- Disponible en sept.-octobre N+2

## ● 3 PRINCIPAUX FICHIERS INDIVIDUELS ESANE (3/3)

### – ECHESA: Contenu des ESA

- Existe depuis 2012 à méthode 2013
- Variables des enquêtes - exemples : Nombre et surface des magasins, Dépenses professionnelles, Sous-traitance...
- Unités échantillonnées à l'exhaustif ou répondantes à l'ESA
- Niveau de diffusion : UL et EP : utilisation des indicatrices diff\_ul ou diff\_ep
- 2 jeux de pondération : poids\_apres\_calage\_ul et poids\_apres\_calage\_ep
- Pour un millésime, 2 années : N-1 et N à contour et méthode N
- Disponible en février-mars N+3

## ● PRODUITS DE DIFFUSION ASSOCIÉS

### – FARE et VAC

- Caractéristiques comptables, financières et d'emploi des entreprises : <https://www.insee.fr/fr/statistiques/5427759>
- Chiffres-clés, TBEF : <https://www.insee.fr/fr/outil-interactif/5367857/>
- Insee référence des entreprises : <https://www.insee.fr/fr/statistiques/5759063>

### – ECHESA

- Fiches sectorielles : <https://www.insee.fr/fr/statistiques/6042002>



## ● DOCUMENTATION

### – Pour chaque fichier (FARE, VAC, ECHESA)

- Note de mise à disposition par année
- Dictionnaire des variables annualisé
- Note détaillée d'utilisation du fichier Fare (nouveau pour le millésime 2019)
- Note « méthodologique » ESA/EAP
- Documentation associée à chaque produit de diffusion (sur insee.fr)

### – Documentation du système Esane

- Présentation générale d'Esane :  
<https://www.insee.fr/fr/metadonnees/source/serie/s1188>
- Présentation complète - Esane méthode disponible sur insee.fr :  
<https://www.insee.fr/fr/information/3056089>





## ● DOCUMENTATION

### – Passage en entreprise

- Insee référence sur la méthode et le mode d'emploi :  
<https://www.insee.fr/fr/statistiques/4255789?sommaire=4256020>
- Passage en entreprise - impact sur le descriptif de l'économie :  
<https://www.insee.fr/fr/statistiques/1906511?sommaire=1906539>
- Courrier des statistiques N2 – 2019 – Le profilage à l'Insee, une identification plus pertinente des acteurs économiques :  
<https://www.insee.fr/fr/information/4168409?sommaire=4168411>

### – Comme données issue des liasses fiscales,

- Formulaires des liasses par régime : BIC, BRN...
- Notice associée par régime et par millésime



**06**

## Exemples d'études à la Direction des statistiques d'entreprises

---



- L'Insee Références « Les entreprises en France » : un ouvrage annuel offrant une vue structurelle complète de notre système productif.
- Depuis l'édition 2019, c'est le concept d'entreprise au sens de la LME qui est retenu comme le niveau d'observation de référence de l'appareil productif.
- Une vue d'ensemble et une série de fiches sectorielles qui dressent une cartographie de l'appareil productif à partir du concept d'entreprise au sens de la LME.

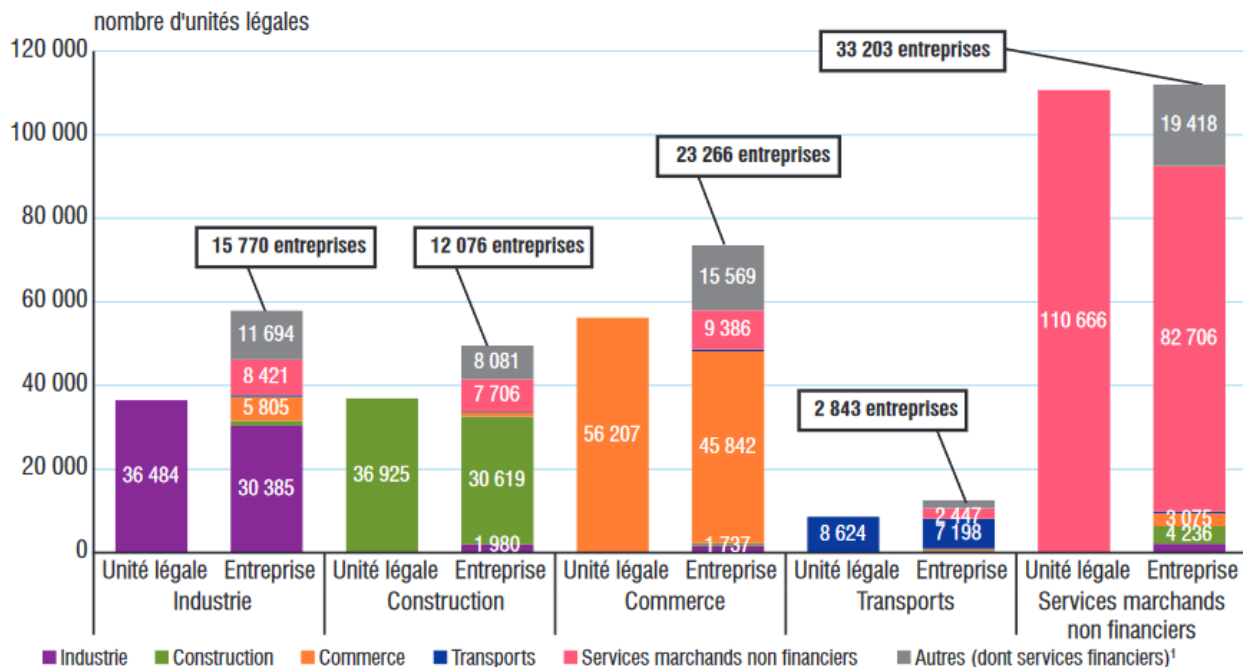


- L'ouvrage aborde de très nombreuses thématiques relatives au système productif : démographie des entreprises, localisation géographique, coût du travail et emploi, investissement, financement, mondialisation, TIC, développement durable, etc.
- L'entreprise au sens de la LME n'est pas l'unique niveau d'observation. Les autres niveaux (« établissement », « unité légale », « groupe ») peuvent avoir chacun leur pertinence en fonction de la dimension du système productif à analyser.
- Le choix du « bon » niveau d'observation, une question délicate ! Quelle articulation entre les différents niveaux d'analyse ?



- La photographie du tissu productif est modifiée à plusieurs titres (cf. Photographie du tissu productif en 2017 », *in Insee Références - Édition 2019*) :
  - Un tissu productif plus concentré ;
  - La dualité entre grandes entreprises et PME est renforcée ;
  - Le poids économique de l'industrie et de la construction est renforcé ;
- Le niveau des principaux agrégats et ratios économiques par secteur est modifié sous le double-effet :
  - Des réallocations sectorielles d'unités légales ;
  - De la consolidation des flux intra-groupes (lorsque les variables comptables ne sont pas additives, comme le chiffre d'affaires par exemple).

## 5. Réallocations d'unités légales entre secteurs d'activité



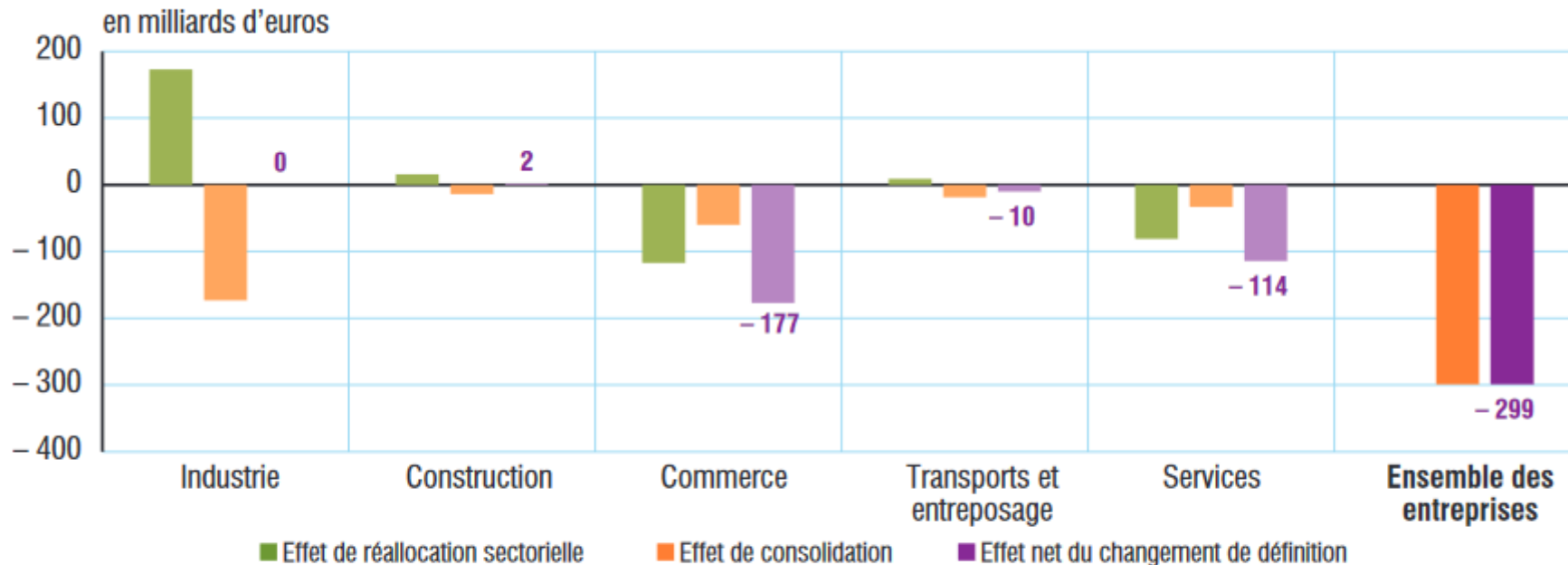
1. Il s'agit d'unités légales qui rentrent dans le champ d'analyse en « entreprise » en tant que sociétés filiales d'entreprises du champ retenu (essentiellement des sociétés *holdings* et des sociétés exerçant des activités auxiliaires de services financiers ou d'assurance).

Lecture : 36 484 unités légales ont une activité principale qui relève de l'industrie. 15 770 entreprises organisées en groupe ont une activité principale qui relève de l'industrie. Ces entreprises sont composées de 57 900 unités légales, dont 30 385 relèvent de l'industrie, 1 176 de la construction, 5 805 du commerce, 8 421 des services marchands non financiers et 11 694 d'autres secteurs (principalement des services financiers).

Champ : France, entreprises constituées en groupe de sociétés, des secteurs principalement marchands non agricoles et non financiers, hors micro-entrepreneurs et régime fiscal de la micro-entreprise.

Source : Insee, *Ésane (données individuelles)*.

## 7. Chiffre d'affaires selon les secteurs : effet de réallocation sectorielle et effet de consolidation



Lecture : la réallocation d'unités légales non industrielles vers les entreprises industrielle accroît le chiffre d'affaires de l'industrie de 173 milliards d'euros par rapport à ce qu'il était en unités légales. Sa consolidation le réduit de 173 milliards d'euros. Le passage en « entreprise » n'a donc pas d'impact sur le chiffre d'affaires de l'industrie.

Champ : France, entreprises des secteurs principalement marchands non agricoles et non financiers, hors micro-entrepreneurs et régime fiscal de la micro-entreprise.

Source : Insee, *Ésane (données individuelles)*.

- Les indicateurs de performance mesurés au niveau de « l'entreprise » sont moins dispersés que ceux mesurés au niveau de « l'unité légale » ;
- Quel impact sur la mesure du niveau de la **productivité du travail** ? (cf. Productivité du travail et du capital : une mesure renouvelée au niveau de l'entreprise », *in Insee Références - Édition 2019* :
  - Rehaussement de la productivité des grandes entreprises ;
  - Rehaussement de la productivité de l'industrie et de la construction ;
  - Moindre dispersion des productivités au sein des secteurs ;
  - Rehaussement de la productivité des entreprises « à la frontière technologique » ;
- Quid de l'impact sur **l'évolution de la productivité** ? Des analyse sont en cours, mais des difficultés apparaissent pour suivre les entreprises LME dans le temps (concept récent et méthodologie de construction complexe). Suivi longitudinal des groupes de sociétés, une bonne alternative ?



**06bis**

## Exemples d'études au Département des Etudes économiques

---



## Étude de l'évolution de la productivité entre 2000 et 2015

- **Données : Esane (2008-2015) et Suse (2000-2007)**
- **Productivité du travail = VA/effectifs**
- **« Productivité à la frontière technologique » : productivité moyenne des 10 % des entreprises les plus productives de leur secteur**

**Symétriquement, « productivité du premier décile » : productivité moyenne des 10 % des entreprises les moins productives de leur secteur**

- **Analyse sectorielle : séparément pour industrie manufacturière, services à basse et moyenne technologies et services de haute technologie**



## Étude de l'évolution de la productivité entre 2000 et 2015 (suite)

- **Analyse de la convergence de la productivité des entreprises vers cette « frontière »**

en modélisant le lien entre gain de productivité en  $t$  et écart à la frontière en  $t-1$

- **Caractérisation des entreprises à la frontière de productivité. Ajout d'informations des DADS et de LiFi**
- ***Insee Références*, Les entreprises en France, édition 2019, Dossier – La productivité en France de 2000 à 2015 : poursuite du ralentissement et hausse modérée de la dispersion entre entreprises.**

## Évaluations de dispositifs d'aide aux entreprises

- **Données : Esane (et Suse avant 2008), appariées à d'autres sources (LiFi, données spécifiques des dispositifs évalués, etc.)**
- **Caractéristiques des entreprises : VA, chiffre d'affaires, investissement, année de création, effectifs, part du CA à l'exportation, etc.**
- *Insee Références, Les entreprises en France, édition 2018, Dossier – Effet du crédit d'impôt pour la compétitivité et l'emploi sur les prix pratiqués par les entreprises.*
- *Document de travail n° 2021-1, Evaluation du dispositif Jeune entreprise innovante (JEI). Un exemple d'application du modèle d'analyse de sensibilité de Rosenbaum.*
- *Document de travail n° G2019/12, Evaluation du Crédit d'impôt innovation.*

# Merci pour votre écoute !

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



11/03/2022



Mesurer pour comprendre