

Comment passer du concept d'unité légale à la définition économique de l'entreprise dans les statistiques sur la R&D ?

Thomas Balcone

Camille Schweitzer



MINISTÈRE
DE L'ENSEIGNEMENT SUPÉRIEUR,
DE LA RECHERCHE
ET DE L'INNOVATION

DGRI

JMS 2018 - COMMENT PASSER DU CONCEPT D'UNITÉ LÉGALE À LA DÉFINITION ÉCONOMIQUE DE L'ENTREPRISE DANS LES STATISTIQUES SUR LA R&D ?

14/06/2018

1

INTRODUCTION

■ Une enquête historique

- Existe depuis 1963
- Répond à un règlement européen
- Collecte au niveau unité légale
- Variable clé : dépenses intérieures de R&D (DIRD) (Stratégie « Europe 2020 »)

■ Nécessité de passer au niveau entreprise

- Nouvelle définition économique de l'entreprise (LME)
- Négociation du futur règlement européen sur les statistiques d'entreprise (FRIBS)
- Mondialisation : importance croissante des grands groupes

■ Données utilisées

- Données de l'enquête R&D 2015 semi-définitives
- Contour des entreprises profilées 2016
- SIRUS 2015



SOMMAIRE

- 1. Comment obtenir des données au niveau entreprise ?
- 2. Comment obtenir un estimateur de la DIRD totale à partir des entreprises répondantes ?
- 3. Quel estimateur au niveau entreprise retenir ?

1. Comment obtenir des données au niveau entreprise ?



1. COMMENT OBTENIR DES DONNÉES AU NIVEAU ENTREPRISE ?

■ Collecte maintenue au niveau unité légale

- Plus simple pour le suivi et le dialogue avec les correspondants
- Analyse intéressante (organisation de la R&D au sein d'une entreprise de plusieurs UL)

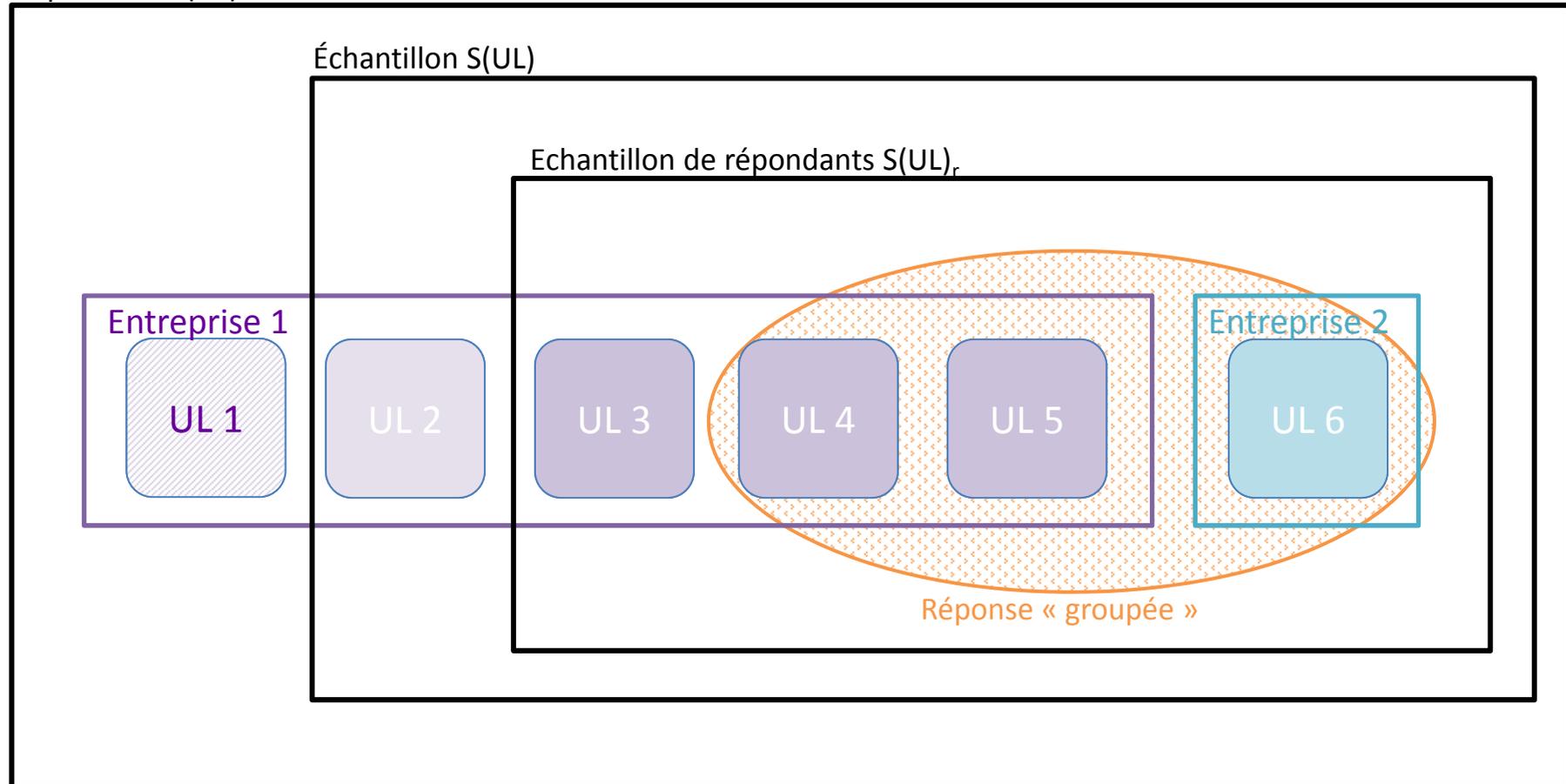
■ Reconstruction de l'information au niveau entreprise

Somme sur les unités légales, mais deux problèmes se posent :

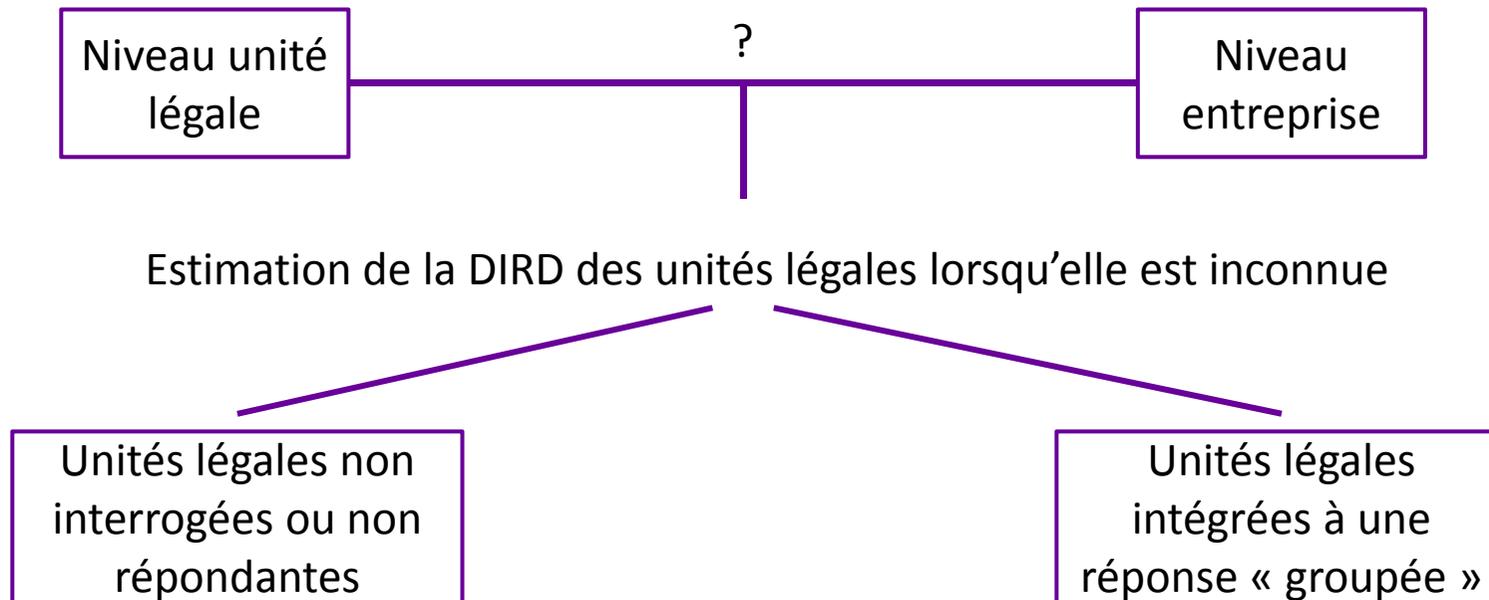
- Données non disponibles pour certaines unités légales
- Existence de réponses « groupées » mêlant plusieurs unités légales

1. COMMENT OBTENIR DES DONNÉES AU NIVEAU ENTREPRISE ?

Population U(UL)



1. COMMENT OBTENIR DES DONNÉES AU NIVEAU ENTREPRISE ?



1. Régression logistique : $DIRD > 0$ ou $DIRD=0$
2. Si $DIRD > 0$, quel montant ?
 - imputation déterministe
 - ou régression linéaire

1. Régression linéaire : quelle DIRD pour chaque UL ?
2. Calcul de la part de chaque UL dans le total estimé sur la réponse
3. Application de cette part au total déclaré

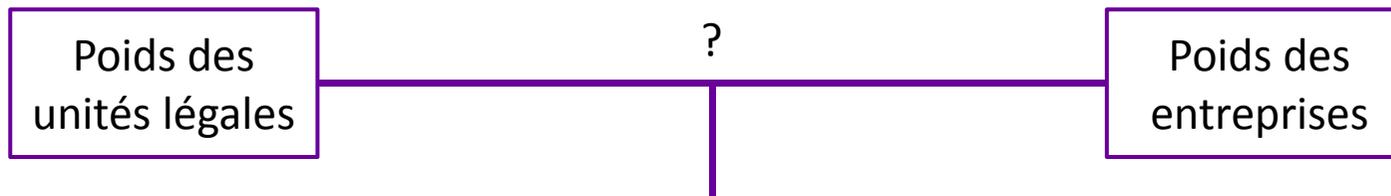
2. Comment obtenir un estimateur de la DIRD totale à partir des entreprises répondantes ?

2. COMMENT OBTENIR UN ESTIMATEUR DE LA DIRD TOTALE À PARTIR DES ENTREPRISES RÉPONDANTES ?

■ Les unités légales des entreprises répondantes

	Nombre d'entreprises répondantes	Nombre d'unités légales (DIRD en %)		
DIRD collectée pour toutes les unités légales de l'entreprise	87 %	64 % (DIRD : 51 %)		
DIRD estimée pour au moins une unité légale de l'entreprise	13 %	DIRD collectée	18 % (DIRD : 37 %)	
		DIRD estimée	Réponses « groupées » relatives à plusieurs entreprises	1 % (DIRD : 11 %)
			Unités légales non interrogées ou non répondantes	17 % (DIRD : 1 %)
TOTAL	S(EP)_r : 8 855	12 695 (DIRD = 30,3 Md€)		

2. COMMENT OBTENIR UN ESTIMATEUR DE LA DIRD TOTALE À PARTIR DES ENTREPRISES RÉPONDANTES ?



Méthode généralisée de partage des poids (MGPP)

$$pond_EP_i^{MGPP} = \sum_{UL_{k,i} \in EP_i \cap S(UL)_r} \theta_{k,i} \times pond_UL_{k,i}$$

MGPP avec liens classiques :
poids basé sur le nombre
d'unités légales dans la
population

**MGPP avec liens
pondérés par la
DIRD**

2. COMMENT OBTENIR UN ESTIMATEUR DE LA DIRD TOTALE À PARTIR DES ENTREPRISES RÉPONDANTES ?

- Exemple d'application de la MGPP avec les deux types de liens

	UL1 $\in S(UL)_r$	UL2 $\notin S(UL)_r$
Poids	4	
DIRD déclarée	100	
DIRD imputée		50
$\widehat{DIRD}_{tot_{UL}}$	400	

MGPP avec liens classiques :

$$\text{Pond}_{EP} = \frac{4}{2 UL} = 2$$

$$\text{DIRD}_{EP} = 150$$

$$\widehat{DIRD}_{tot_{MGPP}} = 150 \times 2 = 300$$

MGPP avec liens pondérés par la DIRD :

$$\text{Pond}_{EP} = \frac{4 \times 100}{150} = 2,67$$

$$\text{DIRD}_{EP} = 150$$

$$\widehat{DIRD}_{tot_{MGPP}} = 150 \times 2,67 = 400$$

3. Quel estimateur au niveau entreprise retenir ?



3. QUEL ESTIMATEUR AU NIVEAU ENTREPRISE RETENIR ?

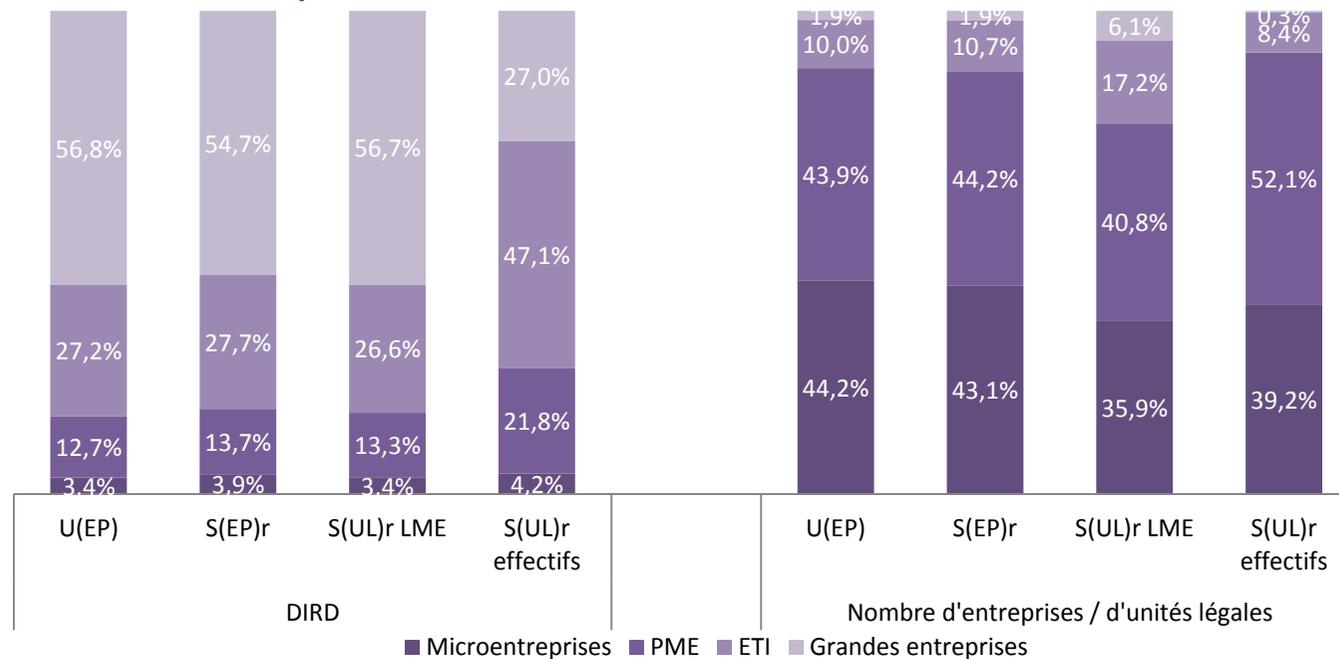
- Écarts relatifs des estimateurs MGPP aux totaux sur U(EP)

	Total sur la population d'entreprises	Écarts relatifs	
		MGPP avec liens classiques	MGPP avec liens pondérés par la DIRD
Nombre d'entreprises	21 466	- 11 %	- 10 %
Nombre d'unités légales	25 965	- 10 %	- 8 %
Chiffre d'affaires (Md€)	2 038	- 3 %	+ 1 %
Effectif (milliers)	5 740	- 6 %	- 4 %
DIRD (M€) (pseudo-total)	31 423	- 2 %	+ 1 %

- L'estimateur MGPP avec liens pondérés par la DIRD semble meilleur

3. QUEL ESTIMATEUR AU NIVEAU ENTREPRISE RETENIR ?

- DIRD et nombre d'unités statistiques par catégories d'entreprise, selon le niveau d'analyse considéré



- Niveau unité légale, définition effectifs : sous-estimation des grandes entreprises
- Niveau unité légale, définition LME : pas d'information sur le nombre d'entreprises

Conclusion et pistes de réflexion



CONCLUSION ET PISTES DE RÉFLEXION

■ Reconstruction des informations au niveau entreprise

- Pas si facile en raison des unités légales non interrogées ou non répondantes et des réponses « groupées »
- Estimation de la DIRD de ces unités légales par régression et/ou imputation

■ Calcul d'un estimateur au niveau entreprise plus pertinent

- Utilisation de la méthode de partage des poids avec liens pondérés par la DIRD
- Meilleure estimation du poids des grandes entreprises et estimation possible du nombre d'entreprises

■ Pistes de réflexion

- Réalisation de simulations pour s'assurer de la qualité de notre estimateur (calcul du biais)
- Calcul d'un estimateur pour les autres variables clés : personnels de R&D et chercheurs ETP
- Refonte du plan de sondage : sondage par grappes

Merci de votre attention

Contacts :

Thomas Balcone – MESRI-SIES – thomas.balcone@recherche.gouv.fr

Camille Schweitzer – MESRI-SIES – camille.schweitzer@recherche.gouv.fr

site internet

<http://www.enseignementsup-recherche.gouv.fr/>