

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

Thomas Balcone



INTRODUCTION

■ L'enquête R&D auprès des entreprises : une enquête historique

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

■ Nécessité de passer au niveau entreprise

- Nouvelle définition économique de l'entreprise (LME)
- FRIBS : diffusion de données au niveau entreprise exigée a priori à partir de 2021 (enquête 2022)
- Une analyse économique plus pertinente : une meilleure compréhension de l'activité 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 la DIRD au niveau entreprise ?

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

- 3. Quel estimateur de la DIRD totale au niveau entreprise retenir ?

1. Comment obtenir la DIRD au niveau entreprise ?



1. COMMENT OBTENIR LA DIRD AU NIVEAU ENTREPRISE ? (1/4)

■ Choisir l'entreprise comme nouvelle unité de collecte?

2 grands risques identifiés :

■ Risque d'une chute du taux de réponse (actuellement au dessus de 90%)

- Comment être sûr de trouver le bon correspondant au sein de l'entreprise?
- Charge de réponse beaucoup plus conséquente a priori

■ Perte d'informations sans de grands changements dans le questionnaire

- L'analyse des activités de R&D par unité légale impossible
- La connaissance des flux financiers intra-groupe impossible
 - > risque très élevé de ruptures de séries pour les variables « dépenses extérieures de R&D » et « ressources externes reçues pour mener des activités de R&D »

1. COMMENT OBTENIR LA DIRD AU NIVEAU ENTREPRISE ? (2/4)

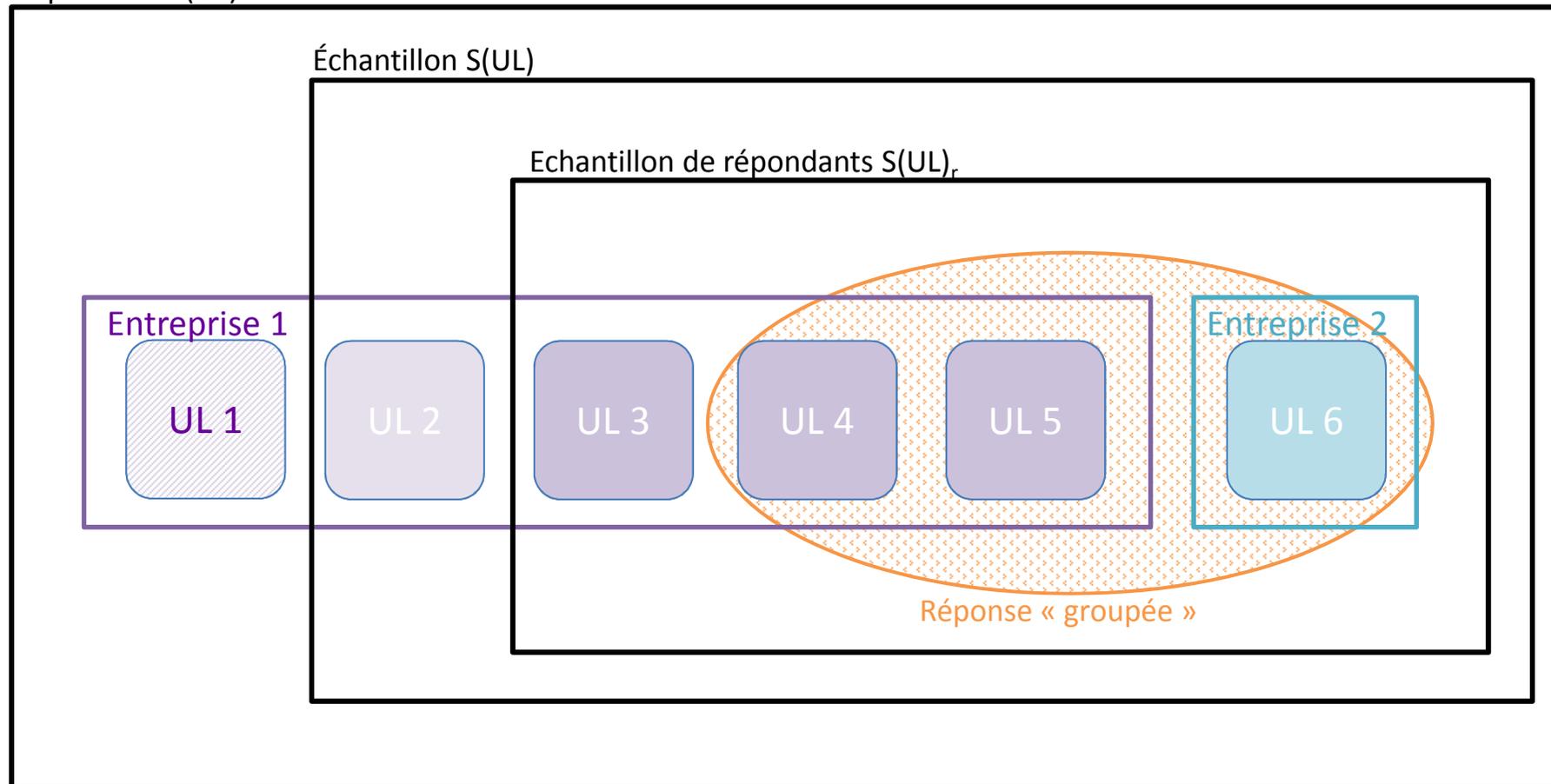
- Collecte maintenue au niveau unité légale (approche moins risquée)

- Reconstruction de l'information au niveau entreprise nécessaire
 - Cas des dépenses intérieures de R&D (DIRD) (variable clé)
 - Variable additive -> Sommer la DIRD des différentes unités légales suffit a priori...
 - ...mais ce n'est pas si simple que cela

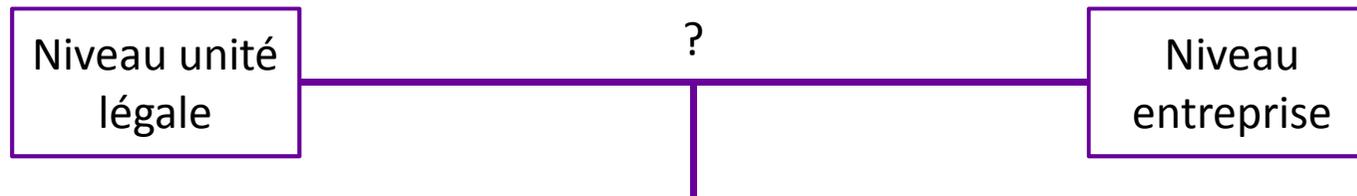
 - Deux problèmes se posent :
 - Données non disponibles pour certaines unités légales au sein d'une entreprise
 - Existence de réponses « groupées » mêlant plusieurs unités légales d'entreprises différentes

1. COMMENT OBTENIR LA DIRD AU NIVEAU ENTREPRISE ? (3/4)

Population U(UL)



1. COMMENT OBTENIR LA DIRD AU NIVEAU ENTREPRISE ? (4/4)



Estimation nécessaire de la DIRD des unités légales lorsqu'elle est inconnue



1. $DIRD > 0$ ou $DIRD=0$?
 - les « grosses » UL : $DIRD > 0$
 - les autres : régression logistique
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 de la réponse « groupée »
3. Application de cette part au total déclaré

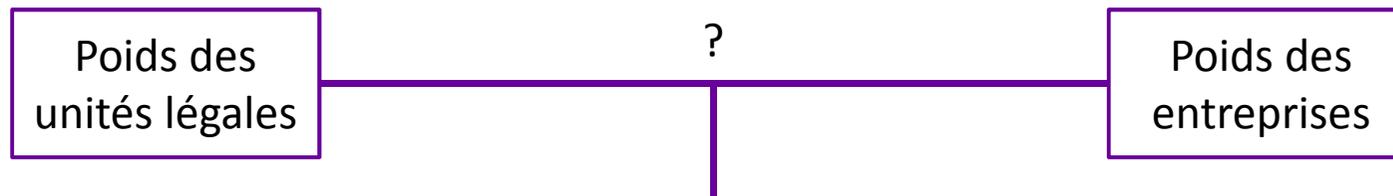
2. Comment obtenir un estimateur de la DIRD totale à partir des entreprises répondantes ?

2. COMMENT OBTENIR UN ESTIMATEUR DE LA DIRD TOTALE À PARTIR DES ENTREPRISES RÉPONDANTES ? (1/3)

■ Les unités légales des entreprises répondantes

	Nombre d'entreprises répondantes	Nombre d'unités légales (DIRD en %)		
Entreprise répondante telle que la DIRD est collectée pour toutes ses unités légales	87 %	64 % (DIRD : 51 %)		
Entreprise répondante telle que la DIRD est estimée pour au moins une de ses unités légales	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 non pondérée = 30,3 Md€)		

2. COMMENT OBTENIR UN ESTIMATEUR DE LA DIRD TOTALE À PARTIR DES ENTREPRISES RÉPONDANTES ? (2/3)



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 ? (3/3)

- 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 de la DIRD totale au niveau entreprise retenir ?

3. QUEL ESTIMATEUR DE LA DIRD TOTALE AU NIVEAU ENTREPRISE RETENIR ? (1/3)

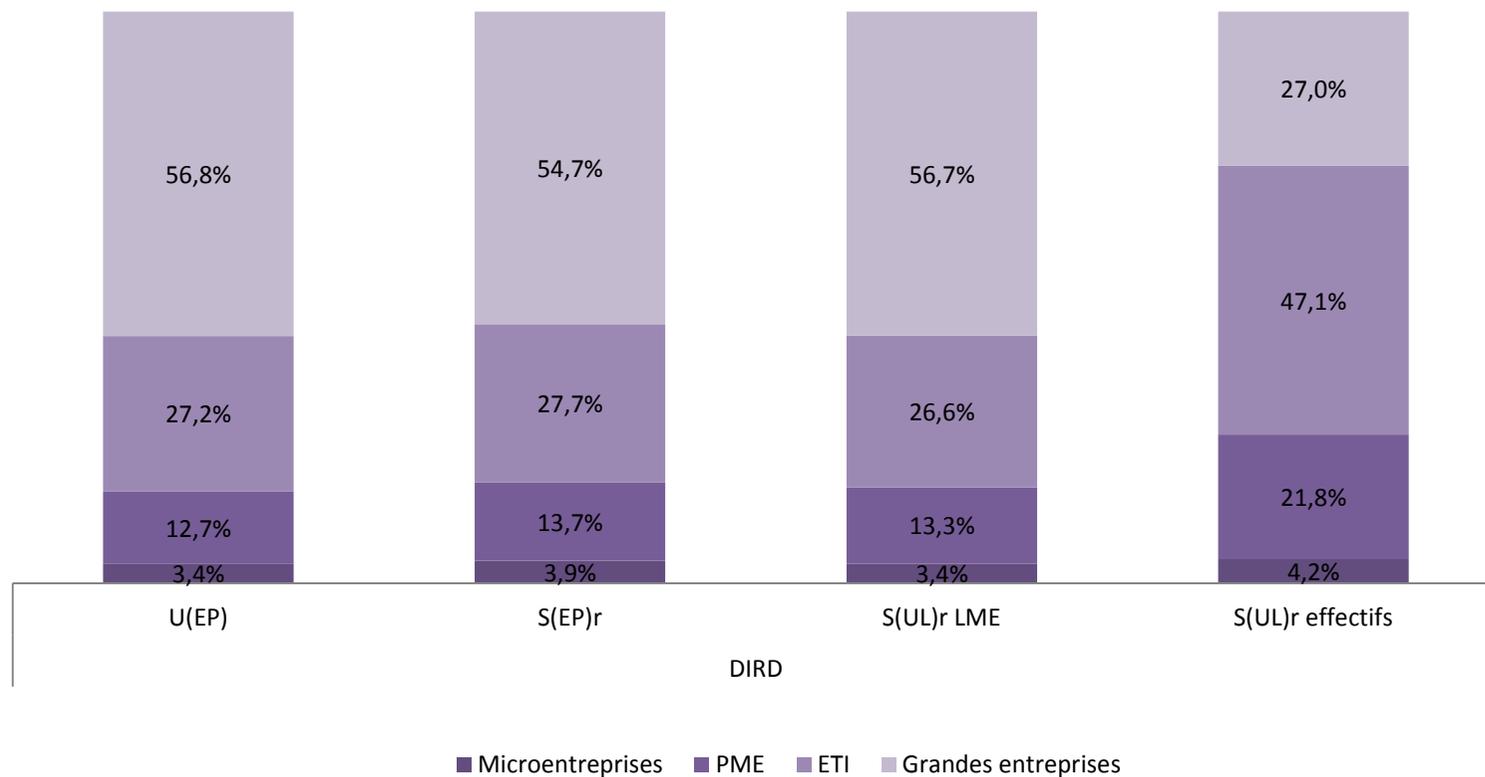
- Écarts relatifs des estimateurs MGPP aux totaux sur U(EP)

	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

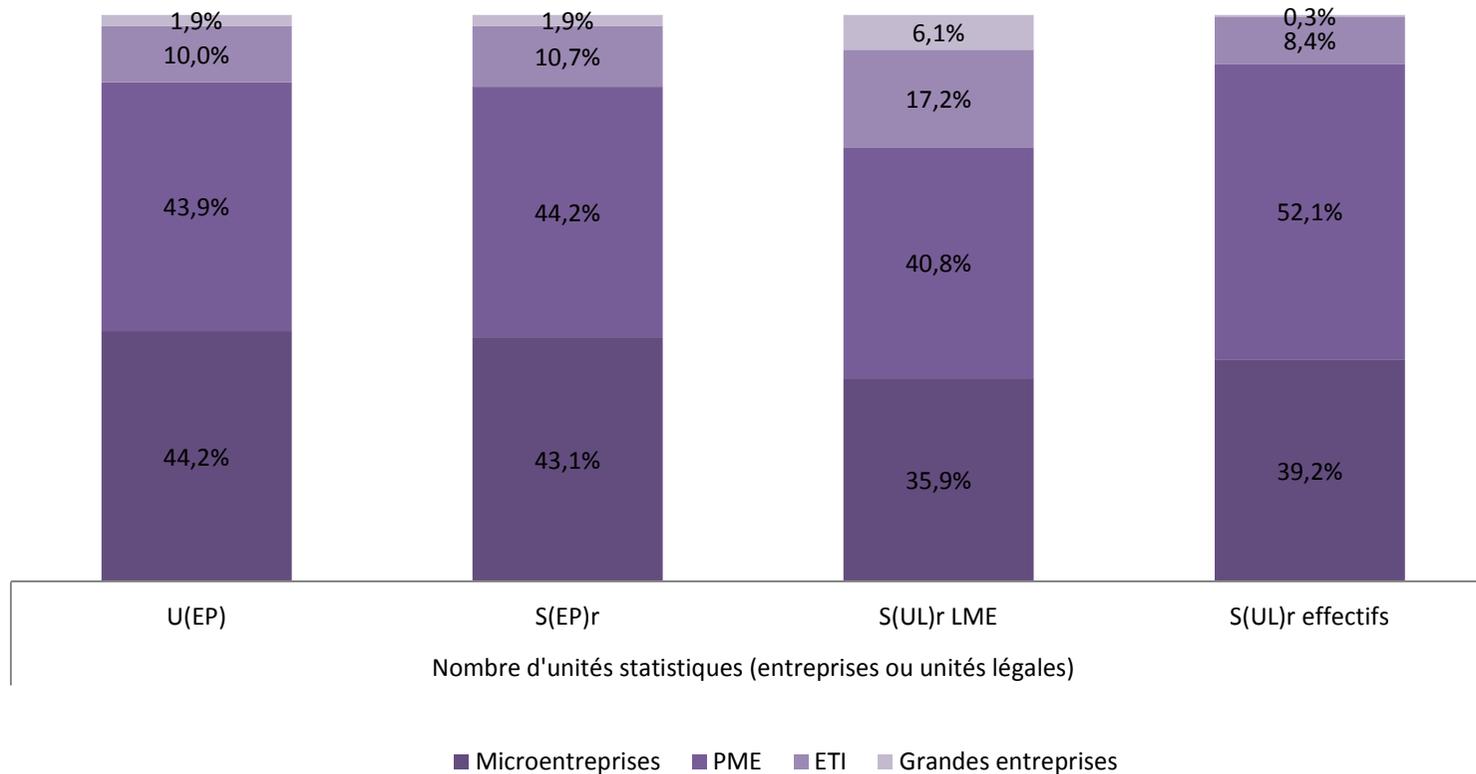
3. QUEL ESTIMATEUR DE LA DIRD TOTALE AU NIVEAU ENTREPRISE RETENIR ? (2/3)

■ DIRD par catégorie d'entreprise



3. QUEL ESTIMATEUR DE LA DIRD TOTALE AU NIVEAU ENTREPRISE RETENIR ? (3/3)

- Nombre d'unités statistiques (entreprises ou unités légales) par catégorie d'entreprise



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 de la DIRD 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 l'estimateur retenu (calcul du biais)
- Calcul d'un estimateur pour les autres variables clés : personnels de R&D et chercheurs ETP
- Choisir l'entreprise comme nouvelle unité de collecte ? Refonte du plan de sondage (sondage par grappes ?) ?

Merci de votre attention

Contact :

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

site internet

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