

Stratégie d'échantillonnage pour les enquêtes couplées

Xème Journées de méthodologie statistique de l'Insee
Sébastien Hallépée, DGCIS

I. Présentation des enquêtes couplées

Qu'est-ce qu'une enquête couplée ?

- Enquête comportant deux volets
 - ✓ Une portant sur les employeurs
 - ✓ Une portant sur les salariés
- Coordination des deux opérations dans la méthodologie et dans l'exploitation
- Les salariés interrogés travaillent dans les entreprises de l'échantillon
- Des exemples récents : « Changements Organisationnels et informatique », Formation Continue, Famille-employeur...

Deux stratégies de tirage possibles

➤ Une stratégie descendante

- ✓ Étape 1 : Sélection d'un échantillon d'entreprises
- ✓ Étape 2 : Sélection d'un échantillon de salariés au sein des entreprises de l'échantillon

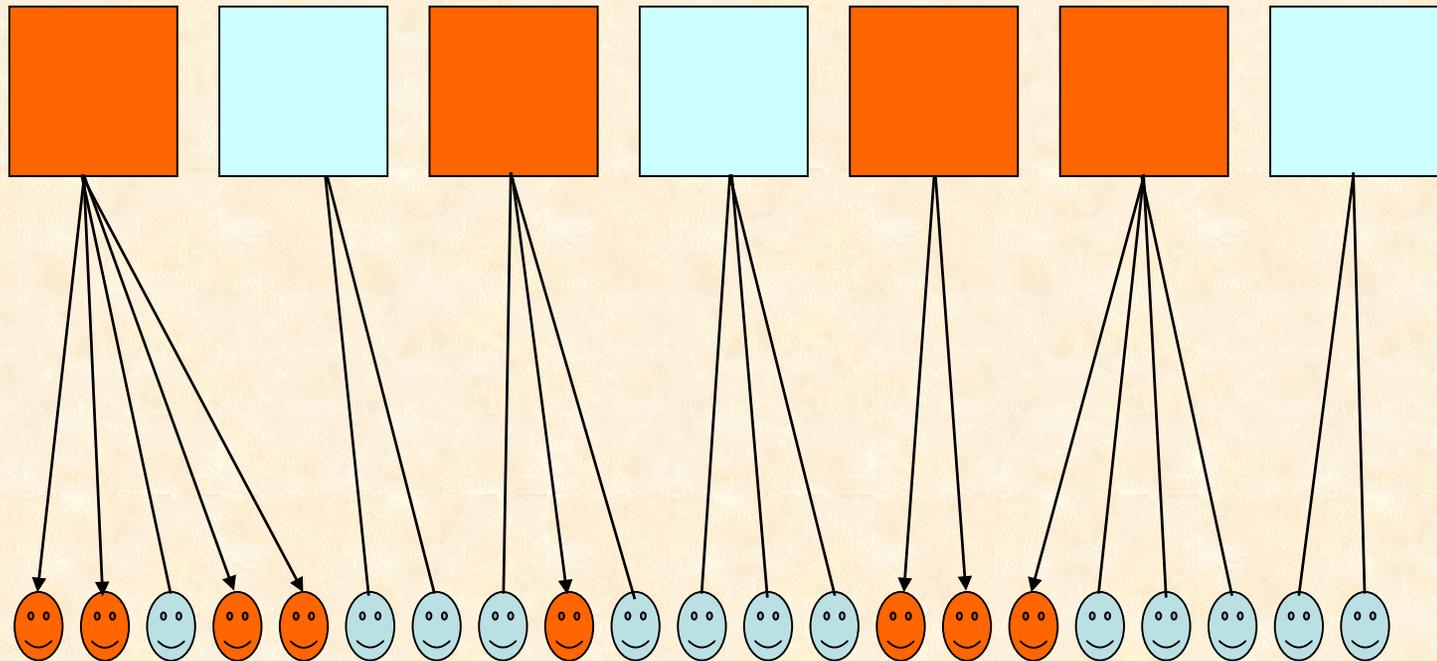
➤ Une stratégie ascendante

- ✓ Étape 1 : Sélection d'un échantillon de salariés
- ✓ Étape 2 : Interrogation de tous les employeurs des salariés de l'échantillon

Schéma de la stratégie descendante, un tirage à deux degrés

Population des entreprises U_1
unités primaires

Échantillon
d'entreprises S_1



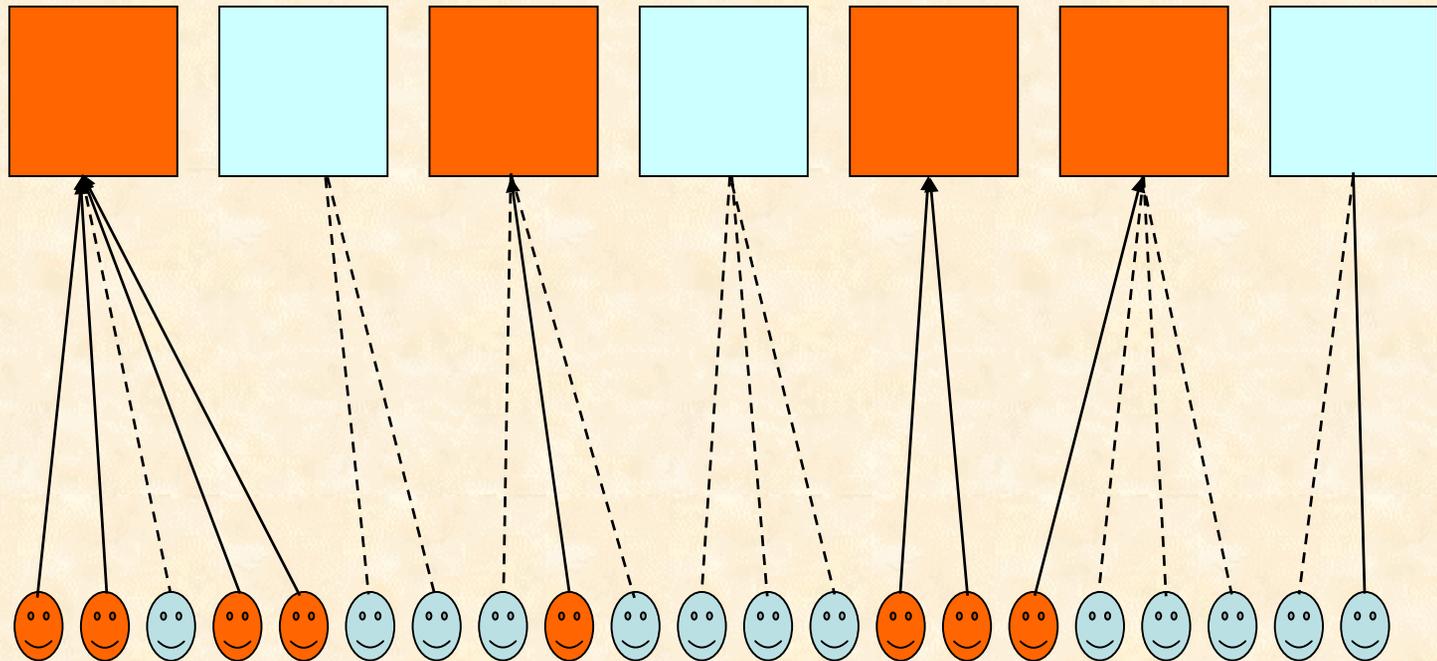
Échantillon de
salariés S_2

Population des salariés U_2
unités secondaires

Schéma de la stratégie ascendante, un tirage indirect des entreprises

Population des entreprises U_2

Échantillon
d'entreprises S_2



Échantillon de
salariés S_1

Population des salariés U_1

population intermédiaire pour les entreprises

Choix de la stratégie

- Stratégie descendante optimale pour l'enquête « entreprises », ascendante pour l'enquête « salariés »
 - ✓ Établir une hiérarchie les priorités menant à un arbitrage
- Contexte peut contraindre la décision
 - ✓ Problème de calendrier
 - ✓ Une des enquêtes est déjà reliée à une enquête existante

II. Stratégie descendante

Stratégie descendante, exemple de l'enquête COI

- Première étape, l'enquête « entreprises »
 - ✓ Tirage stratifié de 17 000 entreprises
 - ✓ Allocation proportionnelle au nombre de salariés
 - ✓ Peu d'influence de l'opération de couplage

- Seconde étape, l'enquête « salariés »
 - ✓ Tirage de 23 000 salariés au sein d'un sous-échantillon d'entreprises répondantes
 - ✓ Nombre de salariés tirés dans une entreprise calculé pour s'approcher d'un tirage auto-pondéré

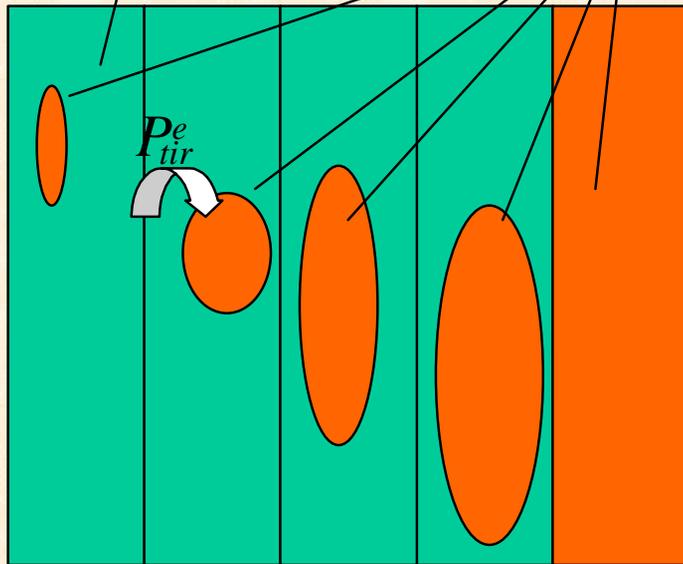
Mode de tirage de l'enquête COI

Étape 1

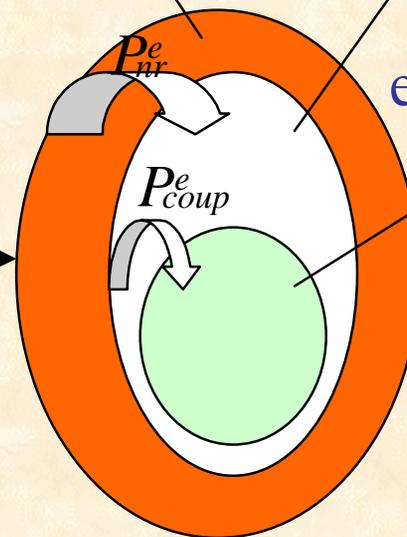
Base de sondage
« entreprises »

Échantillon de l'enquête
« entreprises »

Entreprises
répondantes



Tirage stratifié des
entreprises

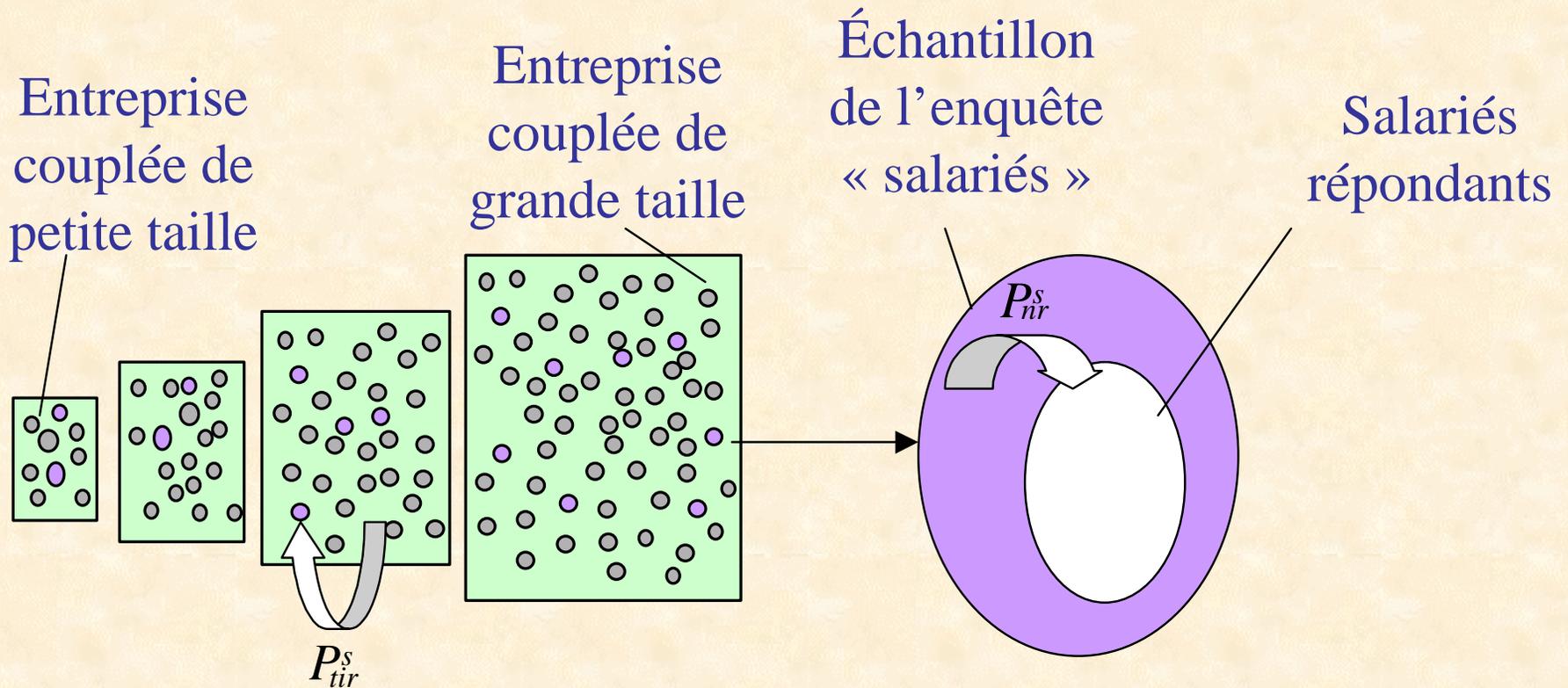


Échantillon des
entreprises couplées

Non réponse et sélection d'un
sous-échantillon pour le couplage

Mode de tirage de l'enquête COI

Étape 2



Tirage et non réponse de l'enquête
auprès des salariés

Influence de la stratégie sur la qualité de l'enquête auprès des salariés

- Information auxiliaire des salariés difficilement mobilisable lors du tirage
- Collecte dispersée géographiquement
 - ✓ 1/3 de l'échantillon dans des zones inhabituelles pour les enquêteurs
- Précision de l'enquête inférieure à un sondage aléatoire simple
 - ✓ Variation d'une variable à l'autre selon l'effet de grappe

Comparaison des précisions des deux stratégies d'échantillonnage

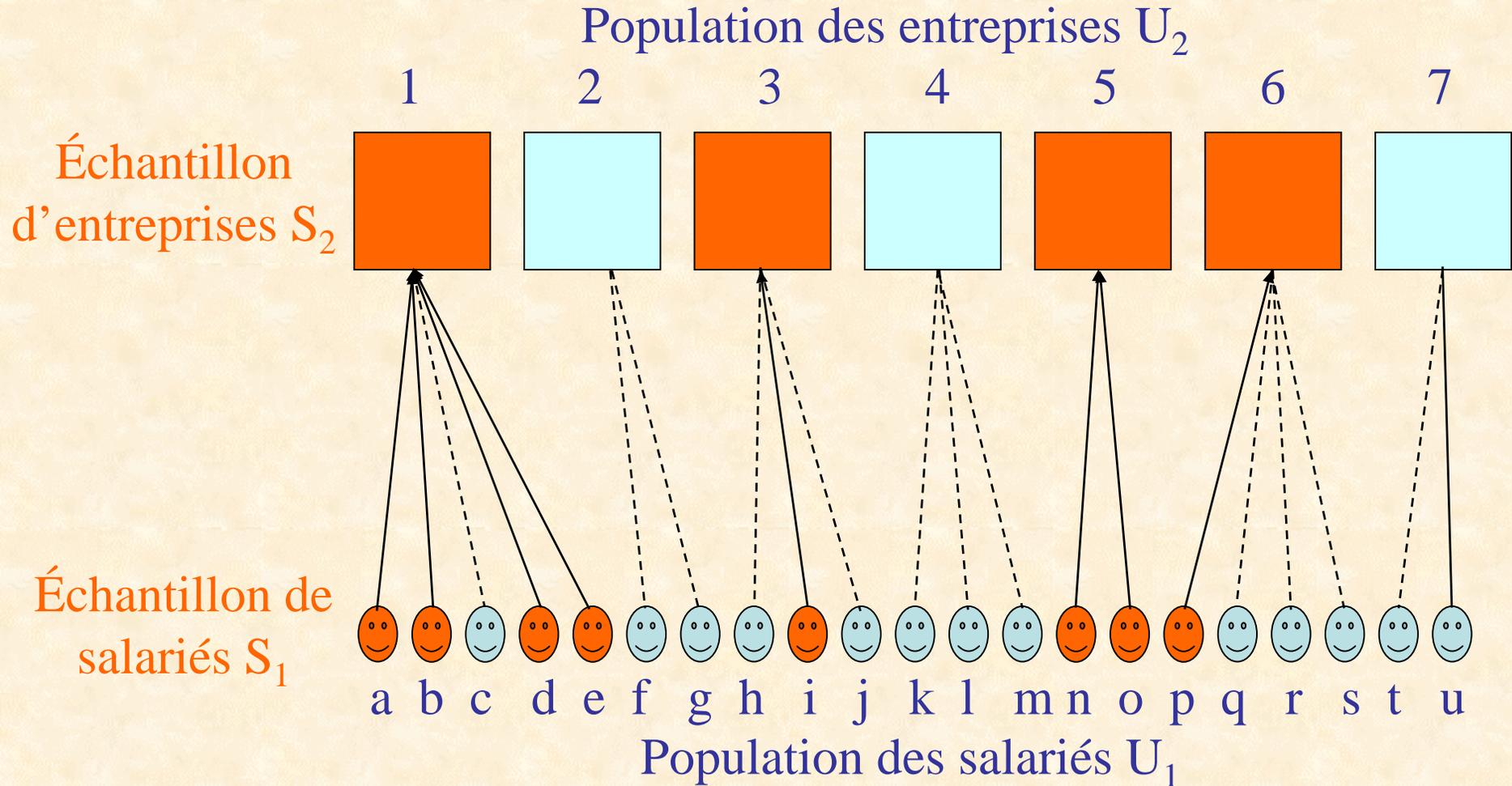
	Proportion de salariés	Coefficient de variation lié à la stratégie descendante (en %)	Coefficient de variation d'un sondage aléatoire simple (en %)	Rapport de précision entre les deux stratégies
Utilisation professionnelle d'un ordinateur fixe	40,9	5,4	1,8	3,0
Utilisation professionnelle d'un ordinateur portable	14,1	6,7	4,0	1,7
Utilisation professionnelle d'un terminal ou d'une console informatique	12,4	6,2	3,2	1,9
Utilisation professionnelle d'autre matériel connecté à Internet ou au SI	17,0	6,2	3,8	1,7
Utilisation professionnelle d'un autre matériel informatique	2,9	7,7	6,4	1,2
Utilise au moins un équipement informatique	51,0	4,6	1,5	3,2
Sympathisant ou adhérent à un syndicat	12,2	13,9	4,8	2,9
Adhérent à un syndicat	11,0	16,4	5,0	3,3

III. Stratégie ascendante

Stratégie ascendante, exemple de l'enquête sur la formation continue

- Première étape
 - ✓ Tirage en grappe de l'échantillon d'individus
 - ✓ Enquête complémentaire à l'enquête emploi
 - ✓ Peu d'influence de l'opération de couplage sur ce volet
- Seconde étape
 - ✓ Interrogation des entreprises où travaillent les salariés de l'échantillon
 - ✓ Tirage par sondage indirect
 - ✓ Un tirage complémentaire
- Collecte des enquêtes en parallèle

Schéma de tirage des enquêtes sur formation continue



population intermédiaire pour les entreprises

Pondération de l'enquête « entreprises »

Application de la méthode du partage des poids

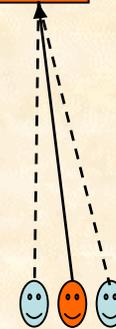
➤ Pondération de l'entreprise j dépend

- ✓ de la (des) pondération(s) du (des) salariés w_i qui ont permis de l'inclure dans l'échantillon
- ✓ de la taille (L_j) de l'entreprise

$$w_{j,PP} = \frac{\sum_{i \in S_1} l_{i,j} \cdot w_i}{L_j}$$

Application

Entreprise 3



h i j

$$w_{3,PP} = \frac{w_i}{3}$$

Influence de la stratégie sur l'enquête auprès des entreprises

- Non prise en compte des caractéristiques des entreprises au moment du tirage
 - ✓ Grandes entreprises potentiellement absentes
 - ✓ Mauvaise représentation sectorielle
- Forte dispersion des pondérations
 - ✓ Pondération inversement proportionnelle à la taille de l'entreprise
 - ✓ Dispersion importante même à taille donnée (influence du tirage en grappe sur les pondérations)
- Précision des estimations inférieure à un tirage stratifié
 - ✓ Meilleure estimation de variables proportionnelles à un nombre de salariés

Comparaison des précisions des deux stratégies d'échantillonnage

	Moyenne	CV associé à la stratégie ascendante (en %)	CV associé à un tirage stratifié (en %)	Rapport de précision entre les deux stratégies
Appartenance à un groupe (en %)	11,1	5,8	3,5	1,7
Entreprise créée après 2000 (en %)	23,5	4,4	2,7	1,6
Pourcentage de cadres	18,3	2,4	1,2	2,0
Entreprises du secteur agricole (en %)	17,9	8,7	0,0	-
Entreprises du secteur de l'énergie (en %)	0,1	55,0	0,0	-
Entreprises du secteur du commerce (en %)	16,6	5,3	0,0	-
Entreprises du secteur des services aux particuliers (en %)	11,1	6,4	0,0	-
Entreprises du secteur de l'éducation, la santé ou l'action sociale (en %)	8,7	6,8	0,0	-

Concilier toutes les approches

Une stratégie mixte

- Possibilité d'utiliser un échantillon complémentaire
 - ✓ Amélioration de la qualité de l'enquête sélectionnée dans un second temps
- Exemple de CVTS
 - ✓ Tirage de toutes les entreprises au delà d'un certain seuil
 - ✓ Gonfler les strates peu représentées suite au tirage indirect