

L'enquête Santé et itinéraire professionnel 2010

*Mise en place d'un double jeu de
pondérations*

Plan

- I. Présentation de l'enquête
 - I. Les objectifs
 - II. Le champ
 - III. La mise en œuvre

- II. Les pondérations longitudinales

- III. Les pondérations transversales

I - Présentation

Les objectifs - Le champ - La mise en œuvre

- L'enquête Santé et itinéraire professionnel (SIP) aborde de façon combinée les questions de conditions de travail et d'emploi et les questions de santé.
- Double interrogation : 2006 et 2010

I - Présentation

Les objectifs - Le champ - La mise en œuvre

- **L'enquête SIP vise un double objectif.**
 - Mieux connaître les déterminants de la santé liés aux grandes caractéristiques des parcours professionnels et des conditions de travail.
 - Repérer en retour l'incidence de l'état de santé sur le parcours professionnel.

- **Objectif supplémentaire ajouté pour 2010 :**
 - Approfondir les connaissances sur l'exposition aux facteurs psychosociaux de risque au travail.



échantillon complémentaire d'actifs occupés

I - Présentation

Les objectifs - Le champ - La mise en œuvre

- Individus âgés de 20 à 74 ans en 2006, vivant en ménage ordinaire en France métropolitaine.
- Pour l'échantillon complémentaire : actifs occupés en 2010, sans critère d'âge.

I - Présentation

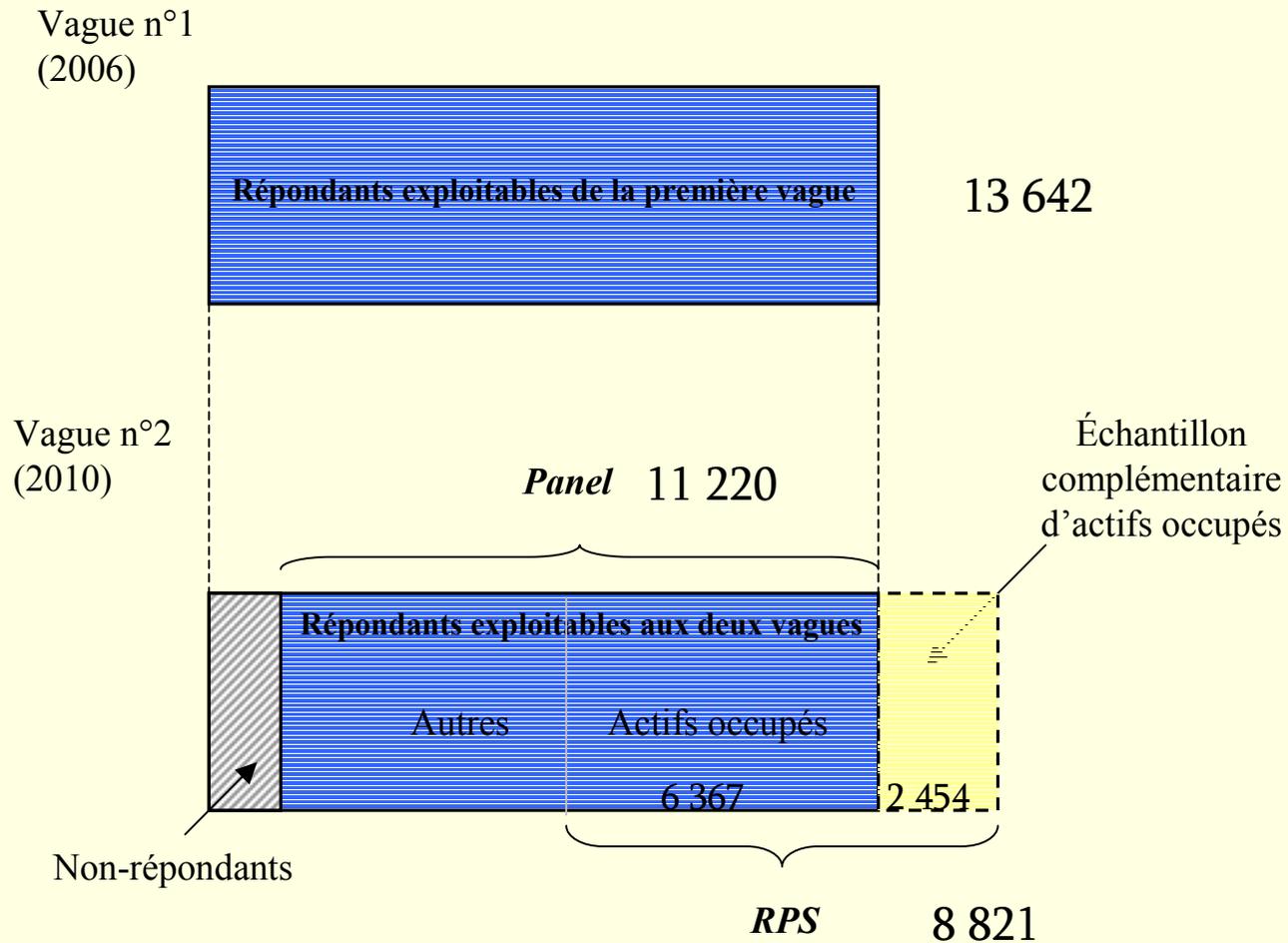
Les objectifs - Le champ - La mise en œuvre

- Les utilisations de la 2nde vague de SIP :
 - **Panel**
 - Répondants aux vagues 2006 et 2010
 - Répondants en 2006, décédés entre les deux vagues
 - **Transversale**
 - Les risques psychosociaux au travail (*RPS*)

⇒ Nécessite la mise en place de deux jeux de pondérations distincts.

I - Présentation

Les objectifs - Le champ - La mise en œuvre

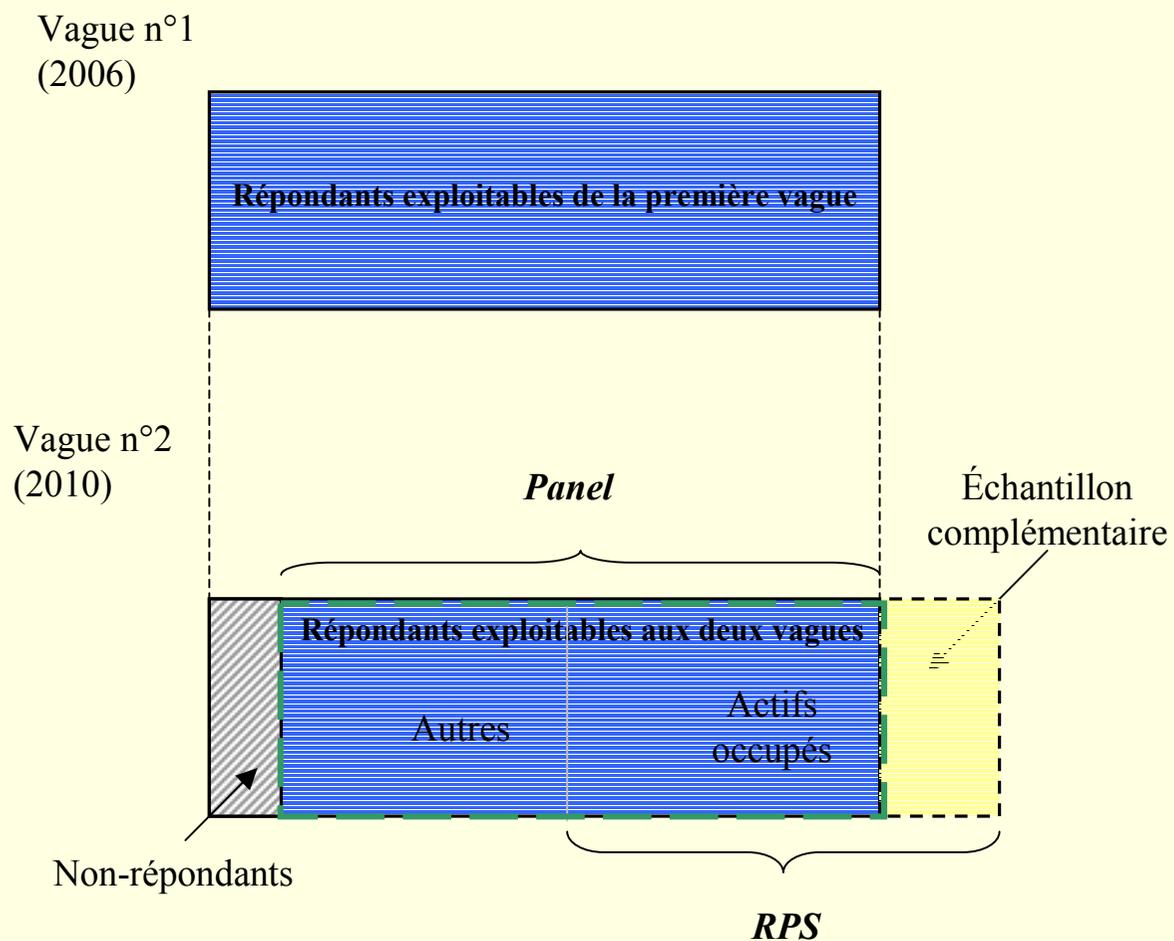


Plan

- I. Présentation de l'enquête
- II. Les pondérations longitudinales
 - I. Traitement de la non-réponse
 - II. Calage
- III. Les pondérations transversales

II – Les pondérations longitudinales

Traitement de la non-réponse - Calage

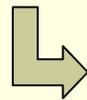


II – Les pondérations longitudinales

Traitement de la non-réponse - Calage

- On dispose des pondérations de la première vague ($W_{v1,i}^{TNR,C}$)

- Correction du biais lié à la non-réponse



création de Groupes de réponse homogène (GRH)

- **Hypothèses :**

- probabilité de réponse constante au sein d'un GRH,
- comportement de réponse/non-réponse indépendant d'un GRH à l'autre.
- Les personnes décédées entre les deux vagues sont exclues de la mise en place des GRH ; elles conservent leur poids initial.

II – Les pondérations longitudinales

Traitement de la non-réponse - Calage

- Variables (issues de la vague 1) utilisées :
 - Situation vis-à-vis de l'emploi
 - Tranche d'unité urbaine
 - L'âge en classes
 - Niveau de diplôme à la fin des études initiales
 - Sexe
 - Santé altérée (état de santé moyen, mauvais ou très mauvais Vs bon ou très bon)

II – Les pondérations longitudinales

Traitement de la non-réponse - Calage

- Méthode utilisée :

- Algorithme CHAID

- 1) L'algorithme regroupe pour chaque variable explicative ses modalités les moins liées au comportement de réponse ;
- 2) Il le fait de proche en proche par paire ou regroupement de modalités ;
- 3) Il retient ensuite la variable explicative dont la significativité corrigée (Bonferroni) est la plus élevée.

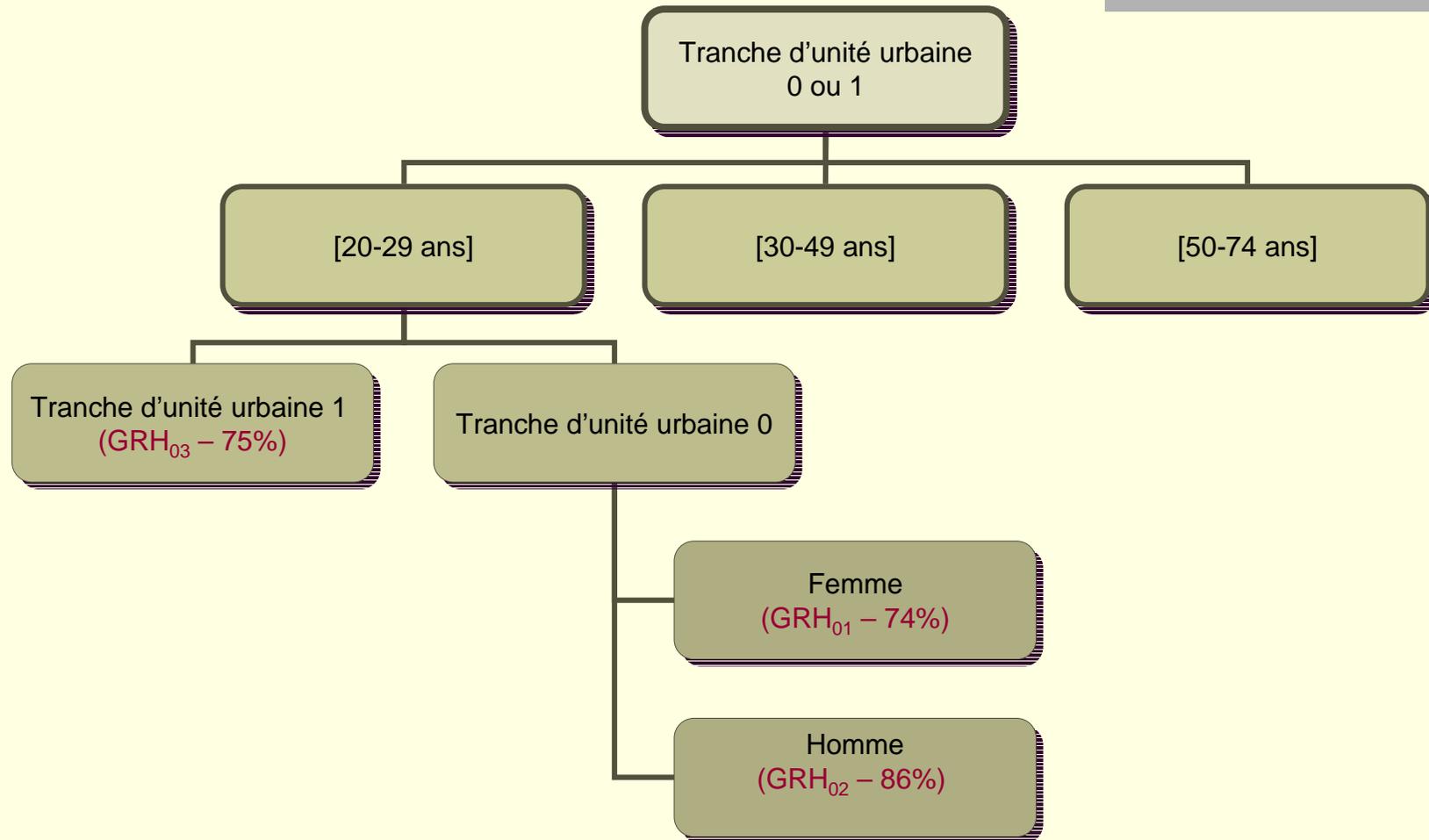
II – Les pondérations longitudinales

Traitement de la non-réponse - Calage

- L'algorithme détermine **27 GRH**.
- Pour chaque GRH h , la probabilité de réponse (P_{Panel}^h) est estimée par le rapport entre le nombre non pondéré de répondants en 2010 et le nombre non pondéré d'individus de cette sous-population répondante en 2006.

II – Les pondérations longitudinales

Traitement de la non-réponse - Calage



II – Les pondérations longitudinales

Traitement de la non-réponse - Calage

- Calage sur six variables :
 - Zone d'études et d'aménagement du territoire
 - Tranche d'unité urbaine
 - L'âge en classes au 31 décembre 2006 par sexe
 - Niveau de diplôme à la fin des études initiales
 - Nationalité de l'enquêté
 - Nombre d'habitants du logement

- Moyenne des quatre enquêtes Emploi trimestrielles de 2006 pour les 20-74 ans.

II – Les pondérations longitudinales

Traitement de la non-réponse - Calage

Quantile	Pondérations initiales $W_{v1,i}^{TNR,C}$	Pondérations traitées de la non réponse $W_{Panel,i}^{TNR}$	Pondérations traitées de la non réponse et calées $W_{Panel,i}^{TNR,C}$	$W_{Panel}^{TNR,C} / W_{Panel}^{TNR}$	$W_{Panel}^{TNR,C} / W_{v1}^{TNR,C}$
Max	21 863	29 319	25 860	1,75	2,68
99%	9 483	12 381	13 403	1,67	2,16
95%	6 072	7 783	8 219	1,55	1,91
90%	4 984	6 222	6 442	1,38	1,75
75%	3 285	4 157	4 348	1,16	1,41
50%	2 479	2 901	2 914	0,96	1,14
25%	1 923	2 321	2 114	0,82	0,96
10%	1 336	1 579	1 589	0,74	0,87
5%	1 210	1 386	1 280	0,71	0,83
1%	891	1 057	943	0,68	0,79
Min	638	705	571	0,66	0,67

II – Les pondérations longitudinales

Traitement de la non-réponse - Calage

- La CSP permet d'apprécier l'impact du calage

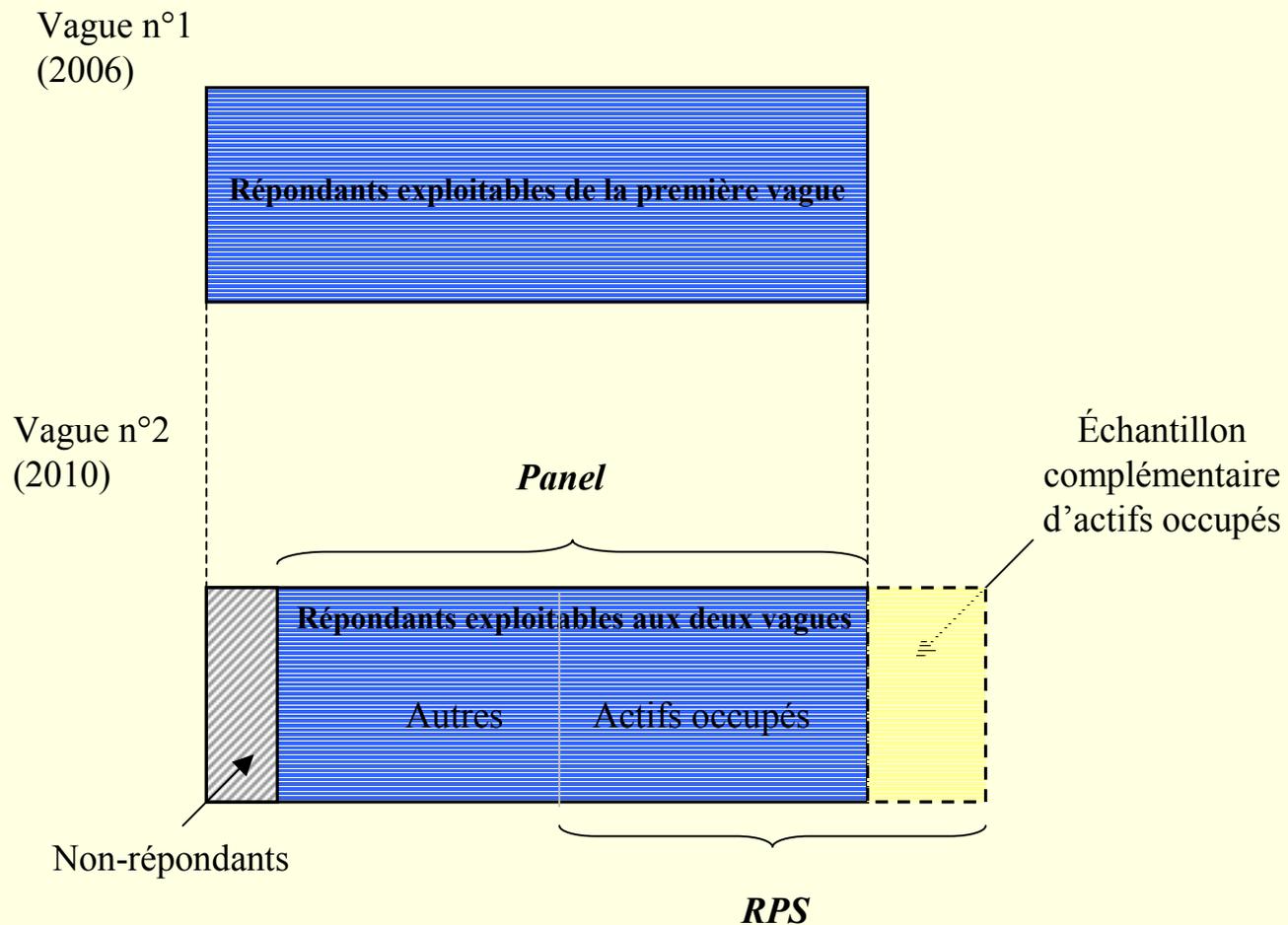
catégories socioprofessionnelles	Pourcentage dans l'échantillon		Pourcentage dans la population
	avant calage	après calage	
0	2%	2%	0%
1	2%	2%	2%
2	4%	4%	4%
3	10%	9%	10%
4	16%	16%	15%
5	20%	20%	20%
6	15%	15%	16%
7	20%	21%	21%
8	10%	11%	13%

Plan

- I. Présentation de l'enquête
- II. Les pondérations longitudinales
- III. Les pondérations transversales
 - I. Traitement de la non-réponse
 - I. Échantillon complémentaire
 - II. Partage des poids
 - II. Calage

III – Les pondérations transversales

Traitement de la non-réponse - Calage



III – Les pondérations transversales

Traitement de la non-réponse - Calage

■ L'échantillon complémentaire

- La probabilité qu'un individu de la population soit tiré dans l'échantillon complémentaire et réponde à l'enquête (P_i) dépend :
 1. de la probabilité initiale de tirage du ménage auquel il appartient (π_m),
 2. de la probabilité estimée de réponse du ménage (P_m),
 3. de la probabilité de tirage de l'individu pour le ménage ($Kish_{m,i}$),
 4. de la probabilité estimée de réponse de l'individu sélectionné sachant qu'il est tiré ($P_{i/m}$).

III – Les pondérations transversales

Traitement de la non-réponse - Calage

■ L'échantillon complémentaire

- La probabilité qu'un individu de la population soit tiré dans l'échantillon complémentaire et réponde à l'enquête (P_i) dépend :

1. de la probabilité initiale de tirage du ménage auquel il appartient (π_m),

- Donnée par le plan de sondage

III – Les pondérations transversales

Traitement de la non-réponse - Calage

■ L'échantillon complémentaire

- La probabilité qu'un individu de la population soit tiré dans l'échantillon complémentaire et réponde à l'enquête (P_i) dépend :
 2. de la probabilité estimée de réponse du ménage (P_m),
 - Construction de **16 GRH** (CHAID) à partir :
 - Tranches d'unité urbaine
 - Taille du ménage en tranches
 - Revenu du ménage en tranches

III – Les pondérations transversales

Traitement de la non-réponse - Calage

■ L'échantillon complémentaire

- La probabilité qu'un individu de la population soit tiré dans l'échantillon complémentaire et réponde à l'enquête (P_i) dépend :
 3. de la probabilité de tirage de l'individu pour le ménage ($Kish_{m,i}$),
 - **Surreprésentation des jeunes actifs occupés de moins de 24 ans** (trois fois plus de chances)

III – Les pondérations transversales

Traitement de la non-réponse - Calage

■ L'échantillon complémentaire

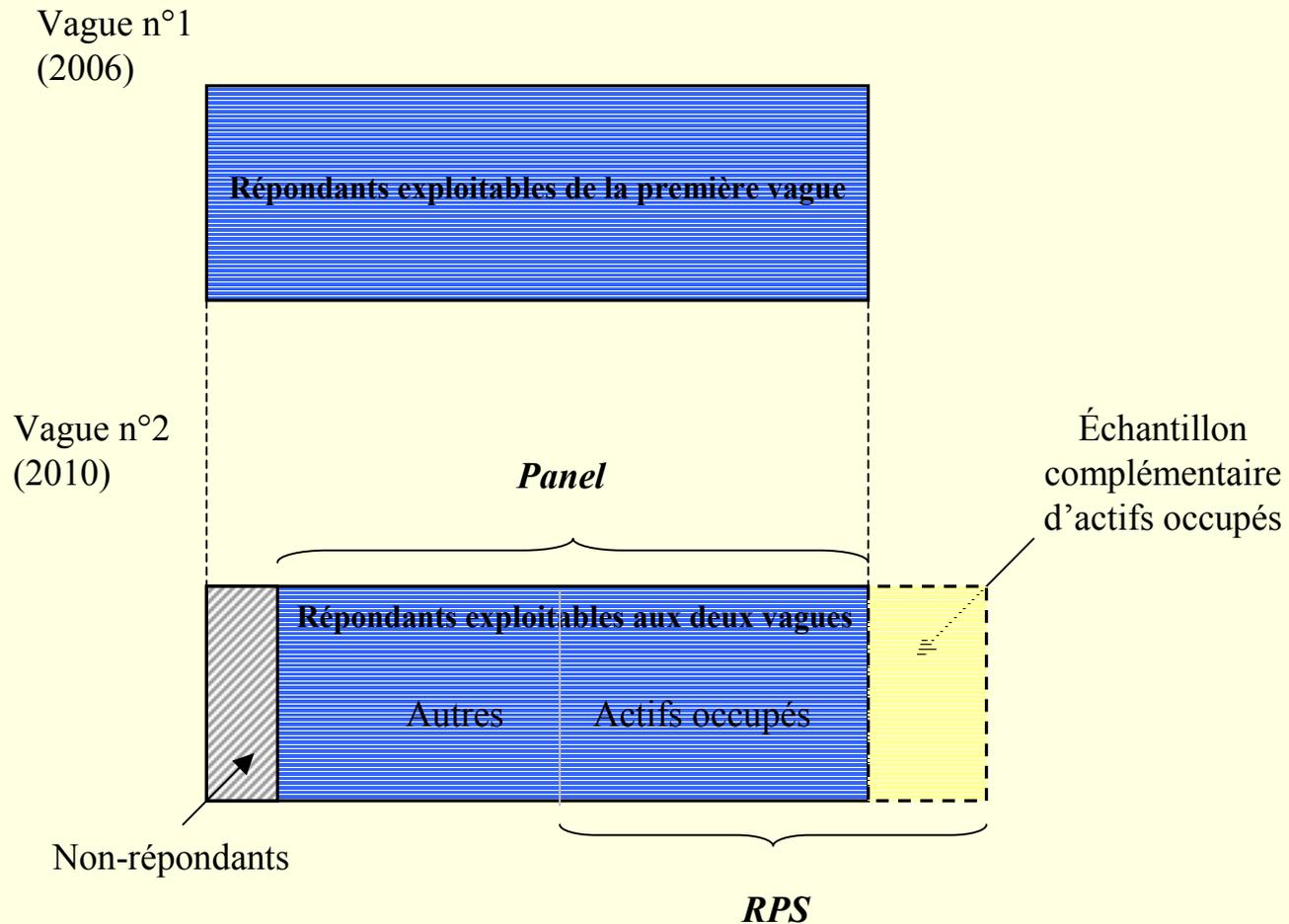
- La probabilité qu'un individu de la population soit tiré dans l'échantillon complémentaire et réponde à l'enquête (P_i) dépend :
 4. de la probabilité estimée de réponse de l'individu sélectionné sachant qu'il est tiré ($P_{i/m}$).



$$P_{i/m} = \frac{2454}{2460} \approx 1$$

III – Les pondérations transversales

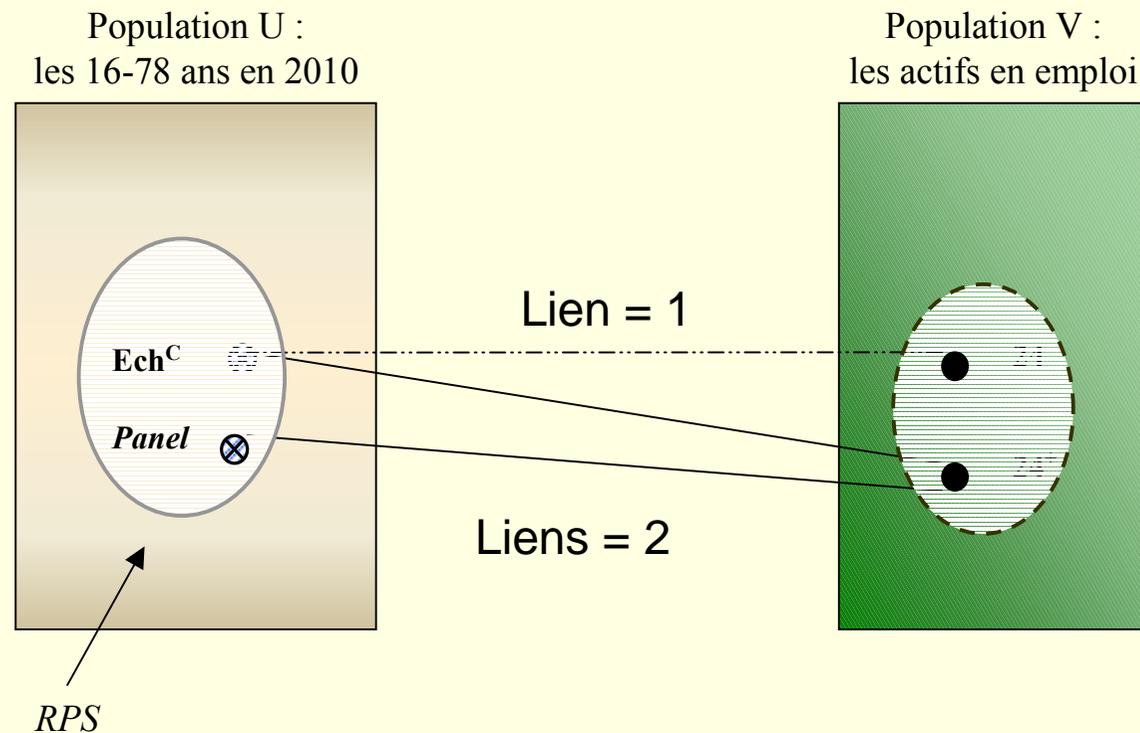
Traitement de la non-réponse - Calage



III – Les pondérations transversales

Traitement de la non-réponse - Calage

- **Le partage des poids :**
 - prise en compte les liens existants entre les deux sous-populations.



III – Les pondérations transversales

Traitement de la non-réponse - Calage

- Pour l'**échantillon complémentaire**, on a

$$\tilde{W}_{ech^C,i}^{TNR} = W_{ech^C,i}^{TNR} \times I_{24}$$

- $I_{24} = 1$ avec si l'enquêté est un jeune actif occupé de moins de 24 ans ;
- $I_{24} = \frac{1}{2}$ si l'enquêté est un actif occupé de 24 ans et plus.

- Pour le **Panel**, on a $\tilde{W}_{Panel,i}^{TNR,C} = \frac{1}{2} \times W_{Panel,i}^{TNR,C}$

III – Les pondérations transversales

Traitement de la non-réponse - Calage

Pondérations avant/après partage des poids

Quantile	Échantillon complémentaire		Actifs occupés du Panel		RPS $W_{RPS,i}^{TNR}$
	avant partage $W_{ech^C,i}^{TNR}$	après partage $\tilde{W}_{ech^C,i}^{TNR}$	avant partage $W_{Panel,i}^{TNR}$	après partage $\tilde{W}_{Panel,i}^{TNR}$	
Max	65 932	38 359	29 319	14 660	38 359
99%	31 958	16 483	13 177	6 588	11 554
95%	18 685	10 204	8 447	4 224	7 989
90%	15 982	8 337	6 801	3 401	5 544
75%	14 224	7 158	4 808	2 404	3 276
50%	7 989	3 995	3 175	1 588	1 890
25%	3 990	2 202	2 317	1 158	1 269
10%	2 298	1 197	1 752	876	987
5%	2 202	1 101	1 441	722	750
1%	1 961	987	1 065	532	550
Min	1 833	917	705	352	352

III – Les pondérations transversales

Traitement de la non-réponse - Calage

- Calage sur quatre variables :
 - Zone d'études et d'aménagement du territoire
 - L'âge en classes au 31 décembre 2010 par sexe
 - Catégorie socioprofessionnelle
 - Nombre d'habitants du logement

- Moyenne des quatre enquêtes Emploi trimestrielles de 2010 pour les actifs occupés de 16-78 ans

III – Les pondérations transversales

Traitement de la non-réponse - Calage

■ Pondérations de la partie RPS

Quantile	Pondérations RPS traitées de la non-réponse $W_{RPS,i}^{TNR}$	Pondérations RPS traitées de la non-réponse et calées $W_{RPS,i}^{TNR,C}$	Rapport des poids
Max	38 359	44 222	1,46
99%	11 554	13 049	1,39
95%	7 989	8 486	1,31
90%	5 544	6 170	1,26
75%	3 276	3 587	1,16
50%	1 890	1 942	1,05
25%	1 260	1 258	0,94
10%	987	938	0,83
5%	759	770	0,78
1%	550	543	0,67
Min	352	215	0,11



**Merci
de votre attention**