

# Biais de rotation de l'enquête Emploi : l'impact des entrées – sorties

*Pascale*  
**BREUIL-GENIER**

*Insee*  
*Direction des statistiques*  
*démographiques et sociales*



*Journées de Méthodologie statistique*

*23 mars 2009*

# Plan

- I. Enquête Emploi et biais de rotation

Présentation, différentes approches méthodologiques et causes possibles

- II. Etude des sorties, entrées, changements de situation d'activité entre deux vagues successives

Une très forte attrition sélective des chômeurs

- III. Simulations et synthèse

Les *variations* de protocole jouent peu, contrairement aux effets de ce protocole sur la situation initiale et les transitions ultérieures.



# I. l'enquête Emploi en continu

- Depuis 2003, une enquête trimestrielle
  - avec collecte en continu
  - dans des délais très contraints
  
- La référence pour le chômage au sens du Bureau International du travail
  
- Echantillon : un panel tournant de logements



# I. Six sous-échantillons interrogés un trimestre donné

|                            |        |        |        |        |        |        |
|----------------------------|--------|--------|--------|--------|--------|--------|
| 1 <sup>er</sup> trim. 2005 |        |        |        |        |        | Visite |
| 2 <sup>e</sup> trim. 2005  |        |        |        |        | Visite |        |
| 3 <sup>e</sup> trim. 2005  |        |        |        | Visite |        |        |
| 4 <sup>e</sup> trim. 2005  |        |        | Visite |        |        |        |
| 1 <sup>er</sup> trim. 2006 |        | Visite |        |        |        |        |
| 2 <sup>e</sup> trim. 2006  | Visite |        |        |        |        | Visite |
| 3 <sup>e</sup> trim. 2006  |        |        |        |        |        |        |
| 4 <sup>e</sup> trim. 2006  |        |        |        |        |        |        |

# I. Le biais de rotation

- Formation « Emploi-Revenus » du CNIS du 8 mars 2007  
« biais de rotation significatif : l'estimation du taux de chômage pour un même trimestre est d'autant plus faible que le rang d'interrogation augmente (à l'exception de la dernière vague) »
- Décision de non-publication des données de l'EEC 2006 et de report du calage des données trimestrielles sur le chômage

Taux de chômage

%

| Rang interrog. | 1    | 2    | 3    | 4   | 5   | 6   |
|----------------|------|------|------|-----|-----|-----|
| Moy. 2005      | 10,6 | 10,1 | 9,8  | 9,8 | 9,0 | 9,7 |
| Moy. 2006      | 10,5 | 10,3 | 10,0 | 9,3 | 9,3 | 9,4 |

# I. Des travaux sur le biais de rotation depuis 2005 au moins

- **D. Goux (notamment JMS 2005)**

En coupe, indices corrigés des fluctuations d'échantillonnage

- **Plan qualité de l'enquête Emploi (2007)**

Analyse longitudinale : comparaison des réponses aux 6 vagues pour un même sous-échantillon

*Taux de chômage (analyse longitudinale)*

%

| <i>Rang d'interrogation du sous-échantillon</i> | 1           | 2    | 3    | 4   | 5          | 6          |
|---|-------------|------|------|-----|------------|------------|
| <i>Ensemble</i>                                 | <b>10,6</b> | 10,4 | 10,1 | 9,9 | <b>9,7</b> | 9,8        |
| <i>Avec statut BIT renseigné 6 fois</i>         | <b>9,0</b>  | 9,1  | 8,8  | 8,8 | 8,8        | <b>8,6</b> |

*Champ : Sous-échantillons entrés entre T1-2003 et T4-2005. Données non pondérées*

# I. De l'analyse longitudinale... à l'analyse des transitions

- L'analyse longitudinale suggère deux types d'effets :
  - les entrées-sorties des répondants contribuent au biais de rotation
  - ceux qui répondent 6 fois répondent différemment selon la vague
- Mais : ces derniers sont très spécifiques et leur 6<sup>ème</sup> réponse n'est pas vraiment comparable à la 1<sup>ère</sup>

*(ils sont moins chômeurs à la 6<sup>ème</sup> vague, ne serait-ce car ils sont plus âgés et n'ont pas déménagé depuis plus d'un an)*

⇒ **Etude des probabilités de changement entre deux vagues successives**

Permet de mieux étudier les différences entre vagues

Décomposition : présence dans le champ, réponse...

Approche descriptive (représentation simplifiée)



# I. Causes possibles des différences entre vagues

- **Différences de protocole délibérées :**
  - 1<sup>ère</sup> et 6<sup>ème</sup> interrogation en visite
  - Protocole simplifié aux vagues intermédiaires (par téléphone)
    - Non-réinterrogation des logements d'inactifs de 65 ans et plus
    - Non vérification du statut des résidences secondaires
    - Pas d'ajout de nouveau logement
  
- **Effets non désirés du protocole :**
  - Probabilité de réponse différentes selon le mode d'interrogation (téléphone/visite), le statut d'activité...
  - Pour les répondants, effets du protocole sur la réponse fournie (effet téléphone, mais aussi proxy, apprentissage...)

# Plan

- I. Enquête Emploi et biais de rotation
- II. Etude des sorties, entrées, changements de situation d'activité entre deux vagues successives
- III. Simulations et synthèse

## II. Des entrées – sorties plus nombreuses que les changements de statut d'activité

| transition                   | effectifs moyen par transition pour un sous-éch. |            |            |            |             |            |
|------------------------------|--|------------|------------|------------|-------------|------------|
|                              | 1 à 2  | 2 à 3      | 3 à 4      | 4 à 5      | 5 à 6       | moy.       |
| tjrs présent – même statut   | 12561  | 12989      | 13010      | 13007      | 12179       | 12749      |
| tjrs présent – autre statut  | <b>646</b>                                       | <b>621</b> | <b>596</b> | <b>588</b> | <b>621</b>  | <b>615</b> |
| tjrs présent – total         | 13208  | 13610      | 13607      | 13595      | 12800       | 13364      |
| va sortir de l'échantillon   | 837  | 877        | 867        | 862        | <b>1493</b> | 987        |
| va entrer dans l'échantillon | <b>1280</b>                                      | 864        | 850        | 697        | <b>1411</b> | 1020       |
| total vague initiale         | 14044  | 14488      | 14474      | 14457      | 14292       | 14351      |
| total vague finale           | 14488  | 14474      | 14457      | 14292      | 14211       | 14384      |

Champ : Sous-échantillons entrés entre T1-2003 et T4-2005. Données non pondérées

## II. Les simplifications de protocole en vague intermédiaire n'ont pas d'effet sur le chômage

- En vague 6 :
  - 261 actifs de 65 ans et plus (10%) ne répondent plus (135 sorties de champ et 126 refus)
  - 107 répondants de logements non réenquêtés avant (tx ch. 15%)  
(prob. d'entrer dans le champ entre deux vagues pour une résidence secondaire : 0,7%)
  - 135 répondants de nouveaux logements en vague 6 (tx ch : 8,7%)

### Impact des simplifications de protocole sur le taux de chômage (pts)

| Rang interrog. | 1 | 2    | 3    | 4     | 5     | 6 |
|----------------|---|------|------|-------|-------|---|
| Impact         | - | 0,00 | 0,00 | +0,01 | +0,02 | - |

Champ : Sous-échantillons entrés entre T1-2003 et T4-2005. Données non pondérées



## II.a Les sorties d'échantillon

- 11,2% des chômeurs ne répondent plus à la vague téléphonique suivante
  - « non-réponse du champ » des chômeurs supérieure de 2 points à celle des actifs occupés en vague téléphonique
  - Un effet visite sur la non-réponse plus dissuasif pour les actifs occupés
- => Un différentiel de non-réponse du champ entre actifs occupés et chômeurs d'un point seulement entre les vagues 5 et 6

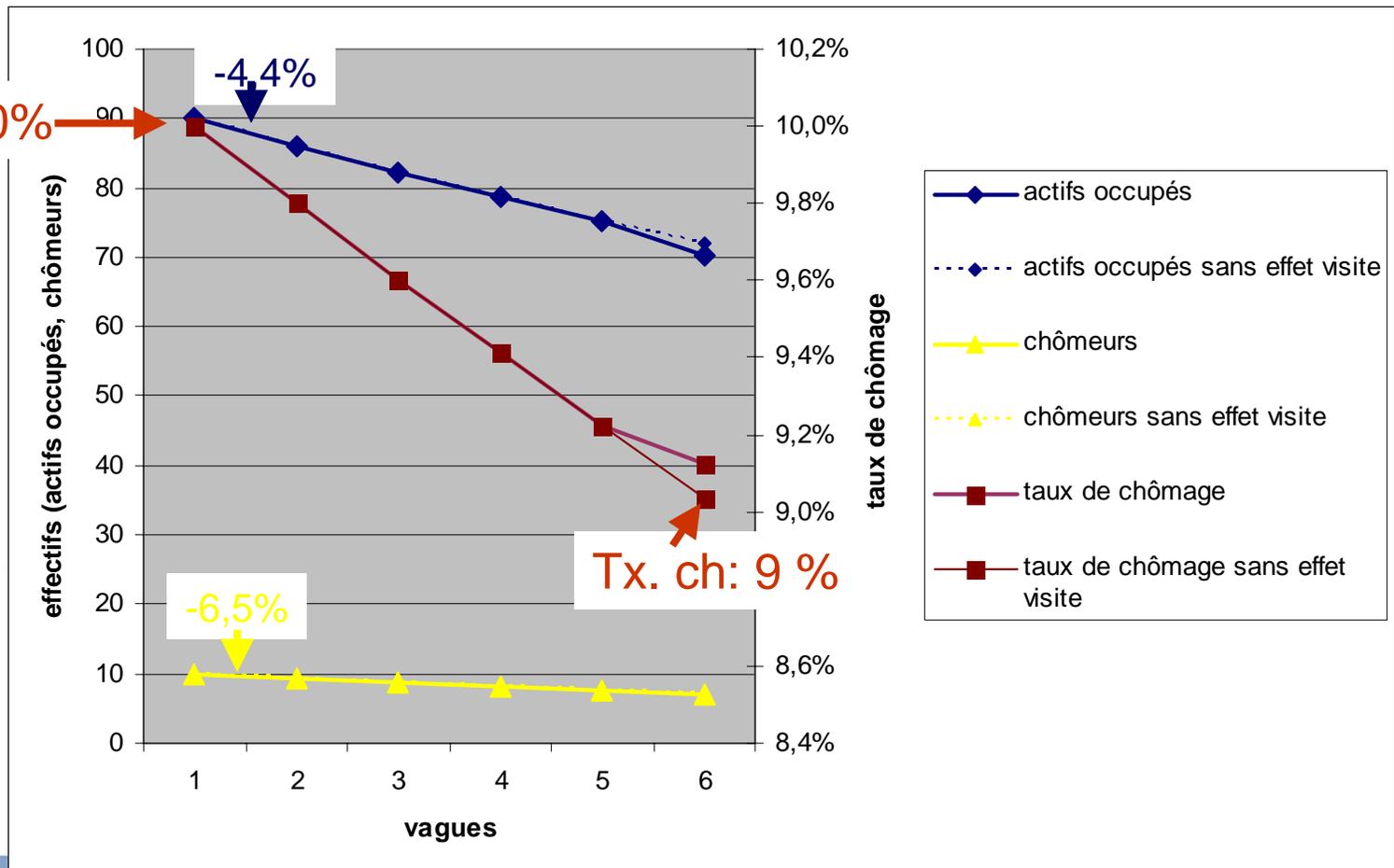
| statut d'activité initial | sortie du champ |                       | non-réponse (du champ) |                    | taux de sortie "réel" estimé (tél.) |
|---------------------------|-----------------|-----------------------|------------------------|--------------------|-------------------------------------|
|                           | probabilité     | effet téléphone (pts) | probabilité tél.       | effet visite (pts) |                                     |
| actif occupé              | 3,0%            | -0,3%                 | <b>4,4%</b>            | <b>2,2%</b>        | 7,2%                                |
| chômeur                   | 5,0%            | -0,2%                 | <b>6,5%</b>            | <b>1,1%</b>        | 11,2%                               |
| inactif                   | 2,4%            | -0,1%                 | 4,3%                   | 1,6%               | 6,6%                                |

*Champ : Sous-échantillons entrés entre T1-2003 et T4-2005. Données non pondérées  
Estimation par maximum de vraisemblance sur les personnes non imputées*



## II.a Le différentiel de taux de réponse (du champ) crée un biais de rotation d'un point !

Simulation sur pop. fictive : impact du seul taux de non-réponse du champ





## II.b Les entrées de logements sans réponse à la vague précédente

- En moyenne par sous-échantillon, 2800 logements sont sans réponse avant transition (57% hors champ et 43% autres non répondants) dont 13% répondent à la vague suivante (17% hors résidences secondaires)
- La probabilité de devenir répondant varie fortement selon le motif d'absence de réponse, est plus forte en vague 2 et en vague 6

| statut initial du logement | présence dans le champ en n+1 |                       | réponse en n+1 (si dans le champ) |                |                    | probabilité entrée totale (tel hors V2) |
|----------------------------|-------------------------------|-----------------------|-----------------------------------|----------------|--------------------|---|
|                            | probabilité                   | effet téléphone (pts) | probabilité téléphone             | effet V2 (pts) | effet visite (pts) |   |
| autre non réponse          | 94,4%                         | 0,5%                  | 21,9%                             | <b>12,0%</b>   | <b>7,2%</b>        | 20,7%                                   |
| hors champ                 | 14,1%                         | -1,0%                 | 59,9%                             | <b>3,5%</b>    | <b>10,2%</b>       | 8,4%                                    |

Champ : Sous-échantillons entrés entre T1-2003 et T4-2005. Données non pondérées

Estimation par maximum de vraisemblance (hors logements neufs, résidences secondaires traitées à part)

## II.b Les enquêtés des logements devenant répondeurs

- Des caractéristiques (nombre par logement, statut d'activité) variant assez peu avec la vague
  - A l'exception des entrants en V2 de logements non-répondeurs du champ en V1 (moins nombreux par logement, taux d'activité plus faible mais de chômage plus fort)
- Mais dépendant un peu plus du motif de non-réponse

|                            | Nombre de répondeurs/<br>logement |          | proportion actifs occupés |          | proportion de chômeurs |          | proportion d'inactifs |          | taux de chômage |          |
|----------------------------|-----------------------------------|----------|---------------------------|----------|------------------------|----------|-----------------------|----------|-----------------|----------|
|                            |                                   | effet V2 |                           | effet V2 |                        | effet V2 |                       | effet V2 |                 | effet V2 |
| statut initial du logement |                                   |          |                           |          |                        |          |                       |          |                 |          |
| autre non réponse          | 2,27                              | -0,17    | 45%                       | -4%      | 7%                     | -1%      | 49%                   | 5%       | 12,9%           | -1,7%    |
| hors champ                 | 2,02                              | .        | 44%                       | .        | 8%                     | .        | 48%                   | .        | 15,4%           | .        |

Champ : Sous-échantillons entrés entre T1-2003 et T4-2005. Données non pondérées

Estimation par maximum de vraisemblance (hors logements neufs, résidences secondaires traitées à part)

Le biais de rotation dans l'enquête Emploi : impact des entrées / sorties - 23 mars 2009



## II.c Les changements d'activité des présents-présents

- Le nombre de transitions est plus élevé entre une vague en visite et une vague en téléphone
  - env. 35 mouvements de plus pour ces transitions (pour 13 000 personnes)
- Les transitions augmentent les effectifs en emploi (+13 par vague et sous échantillon) et diminuent ceux au chômage (-7)
  - effet de questionnaire, protocole, biais de sélection...

| (par rapport au nb total d'observations initiales) | 1 à 2 | 2 à 3 | 3 à 4 | 4 à 5 | 5 à 6 | Moy.  |
|--|-------|-------|-------|-------|-------|-------|
| <i>taux de changement</i>                          | 4,9   | 4,6   | 4,4   | 4,3   | 4,9   | 4,6   |
| actif occupé                                       | 0,20  | 0,10  | 0,06  | 0,02  | 0,09  | 0,09  |
| chômeur  | 0     | -0,11 | -0,05 | 0     | -0,08 | -0,05 |
| inactif  | -0,20 | 0,02  | -0,02 | -0,02 | -0,01 | -0,05 |

Champ : Sous-échantillons entrés entre T1-2003 et T4-2005. répondants aux deux vagues de la transition.

Données non pondérées





# Plan

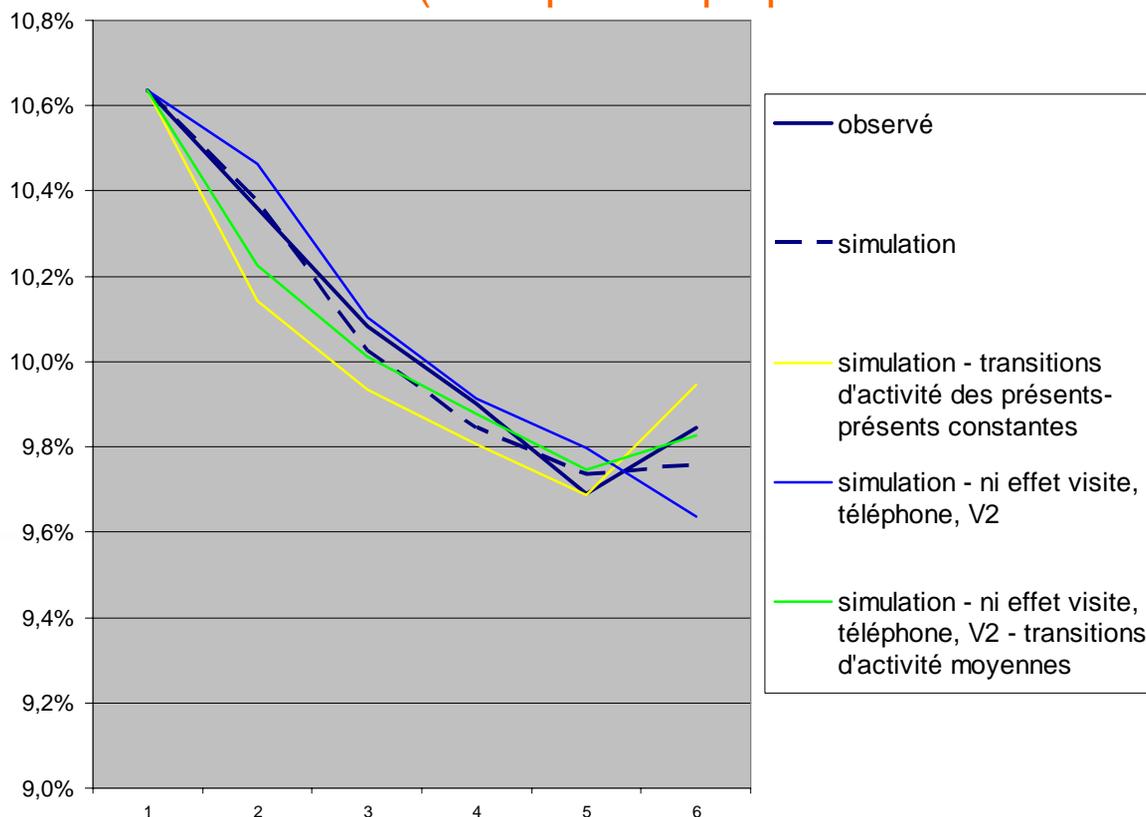
- I. Enquête Emploi et biais de rotation
- II. Etude des sorties, entrées, changements de situation d'activité entre deux vagues successives
- III. Simulations et synthèse

# III.a Synthèse : effet des variations de protocole

Simulation des effectifs par statut en fonction des paramètres estimés pour les sorties, entrées, transitions présents-présents

=> Supprimer les variations de ces paramètres en fonction des vagues ne supprime pas le biais de rotation (mais peut expliquer la remontée en V6)

Taux de chômage



Simulation (hors logements neufs, résidences secondaires, sans protocole simplifié en vague intermédiaire)



## III.b Synthèse : le biais de rotation est avant tout dû aux effets moyens du protocole

Il reflète le fait que la structure par activité à t n'est pas cohérente avec les probabilités de transitions observées

Exemple :

$$P(\text{chômeur en } t+1 \mid \text{chômeur en } t) = 0,5$$

$$P(\text{actif occupé en } t+1 \mid \text{chômeur en } t) = 0,5$$

$$P(\text{chômeur en } t+1 \mid \text{actif occupé en } t) = 0,1$$

$$P(\text{actif occupé en } t+1 \mid \text{actif occupé en } t) = 0,9$$

Si la population de départ correspond à l'équilibre de long terme, pas de biais de rotation →

| équilibre | chômeurs | actifs occupés | taux de chômage |
|-----------|----------|----------------|-----------------|
| t=1       | 50       | 250            | 16,7%           |
| t=2       | 50       | 250            | 16,7%           |
| t=3       | 50       | 250            | 16,7%           |
| t=4       | 50       | 250            | 16,7%           |
| t=5       | 50       | 250            | 16,7%           |
| t=6       | 50       | 250            | 16,7%           |



# III.b Synthèse : le biais de rotation est avant tout dû aux effets moyens du protocole

Si les probabilités de transition observées sont biaisées

$$P(\text{chômeur en } t+1 \mid \text{chôm. en } t) = 0,45$$

$$P(\text{actif occ. en } t+1 \mid \text{chôm. en } t) = 0,45$$

=> Biais de rotation + biais à long terme

| attrition | chômeurs | actifs occupés | taux de chômage |
|-----------|----------|----------------|-----------------|
| t=1       | 50       | 250            | 16,7%           |
| t=2       | 48       | 248            | 16,1%           |
| t=3       | 46       | 244            | 15,9%           |
| t=4       | 45       | 240            | 15,8%           |
| t=5       | 44       | 237            | 15,8%           |
| t=6       | 44       | 233            | 15,8%           |

ou si la population de départ n'est pas représentative

=> Biais de rotation (mais convergence)

| non-repr. | chômeurs | actifs occupés | taux de chômage |
|-----------|----------|----------------|-----------------|
| t=1       | 55       | 245            | 18,3%           |
| t=2       | 52       | 248            | 17,3%           |
| t=3       | 51       | 249            | 16,9%           |
| t=4       | 50       | 250            | 16,8%           |
| t=5       | 50       | 250            | 16,7%           |
| t=6       | 50       | 250            | 16,7%           |



**Merci pour votre attention !**