

L'enquête Emploi en Continu 2019-2028: étude du second degré du plan de sondage

Thomas Sauvaget (SSP75), Thomas Deroyon (Drees), Thomas Merly-Alpa (Ined)



Travaux effectués en 2019 à la Division Sondages (Insee)



Journées de méthodologie statistique de l'Insee

2022

31 mars 2022

Plan

- Contexte et problématique : un nouvel échantillon EEC
- Rappels sur le premier degré du plan de sondage
- Etude du second degré du plan
- Construction de « grappes-univers »
- Tirage des grappes à enquêter
- Conclusion

Tous les fonds de cartes utilisés sont © les contributeurs OpenStreetMap
www.openstreetmap.org

Contexte et Problématique : un nouvel échantillon EEC

- **Les enquêtes ménage à l'Insee : situation avant 2018**

Deux échantillons conçus séparément :

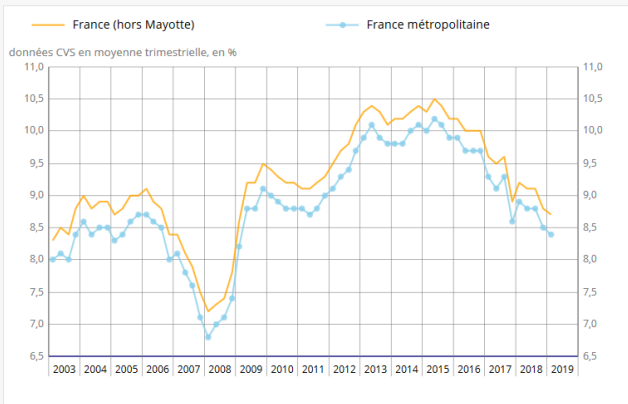
- un échantillon pour l'Enquête Emploi en Continu (EEC)
- un Échantillon-Maître (EM) pour les autres enquêtes ménage

- **En 2018 : une opportunité de coordination**

- l'échantillon 2009 de l'EEC est à renouveler
 - il avait été construit sur la TH (i.e. pas sur le RP)
 - il faudra conserver son organisation rotative (voir plus loin)
- l'EM Octopusse arrive en fin de vie
 - il avait été construit sur le recensement rénové (5 groupes annuels)
 - 10 ans plus tard les groupes RP sont déséquilibrés : une nouvelle méthodologie est nécessaire

- Objectif principal de l'EEC :

Taux de chômage au sens du BIT



Estimation à +/- 0,3 point près du niveau du taux de chômage et de son évolution d'un trimestre à l'autre.

Champ : population des ménages, personnes de 15 ans ou plus.

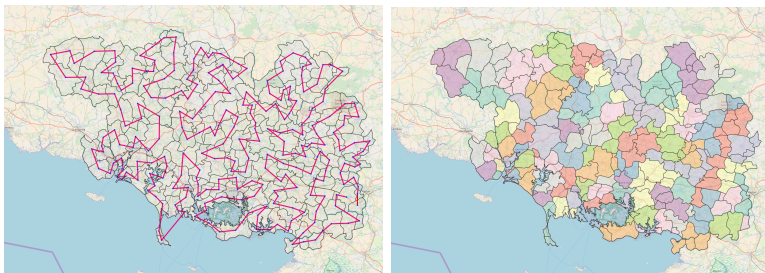
Source : Insee, enquête Emploi.

- Organisation rotative de l'EEC à conserver :
 - enquête trimestrielle
 - échantillon de logements, environ 70 000
 - utiliser la TH (via Fidéli)
 - enquête aérolaire : tirage de groupes de logements proches
 - aires = zones petites d'environ 20 logements
 - panel rotatif + conserver acquis enquêteurs :
 - secteur = groupe de 6 (ou 7) aires proches
 - tirage de secteurs : équilibré selon variables emploi et géographiques

- Durée de vie de l'échantillon :
 - aire = 6 trimestres consécutifs (1,5 an)
 - donc durée de vie des secteurs = $6 \times 1,5 = 9$ ans

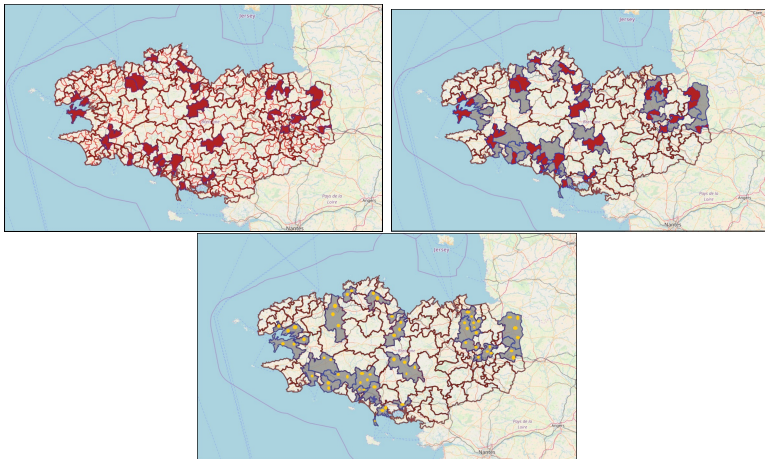
Premier degré du plan de sondage : tirage des secteurs

- Tirage coordonné d'échantillons : EM et EEC
- Méthode retenue : tirage indirect
- étape 0 : découpage métropole en 5064 UP (voyageur du commerce entre communes) : 2 500 logements minimum par UP (250 logements/an sur 10 ans)



À droite : une couleur = une UP = un groupe de communes

- étape 1 : tirage des UP (rouge)
- étape 2 : récupération des UC associées aux UP tirées (gris)
- étape 3 : tirage des secteurs dans les UC récupérées (points jaunes)
→ tirage stratifié et doublement équilibré



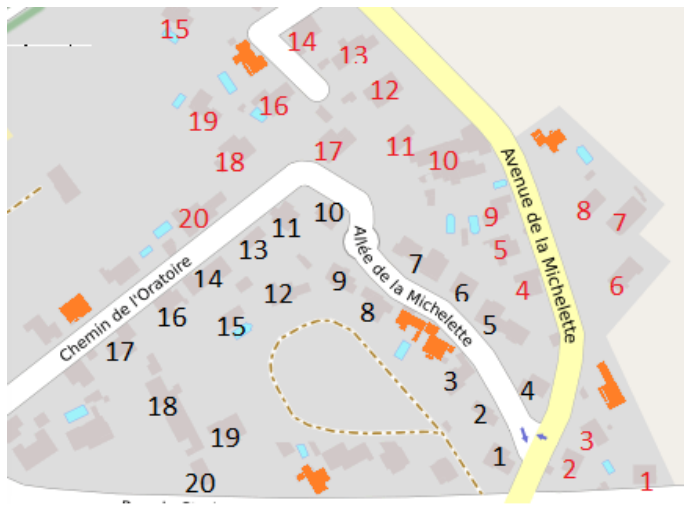
Construction des secteurs :

- étape A : distinction entre RP et RNP



- Résidence Principale (RP)
- Résidence Non Principale (RNP)

- étape B : création de grappes de RP 2016



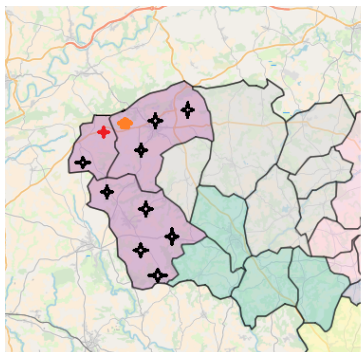
Regroupement des RP en grappes
(environ 20 par grappe)

Étude du second degré du plan de sondage

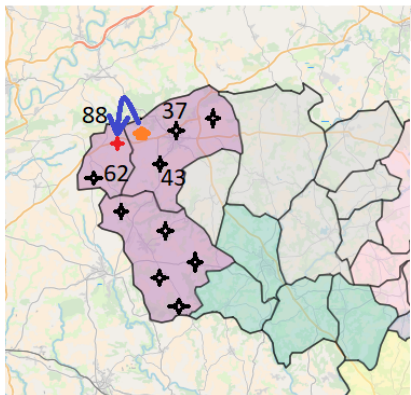
- Objectif : obtenir des grappes à enquêter ayant RP et RNP
- Contraintes à respecter :
 - les logements sont groupés par « étage »(contrainte de collecte) :
 - soit étage d'immeuble, soit maison individuelle
 - maximum de 40 logements par grappe, dont maximum 24 RP
 - gérer le fait que le statut RP/RNP évolue chaque année
 - gérer les logements nouveaux et logements détruits chaque année
- Solution :
 - rattacher les RNP 2016 et les logements nouveaux 2017 et 2018 aux secteurs, tout en respectant les étages
→ obtention de « grappes-univers » 2018
 - y réaliser un tirage de second degré stratifié à 2 phases

Construction de « grappes-univers »

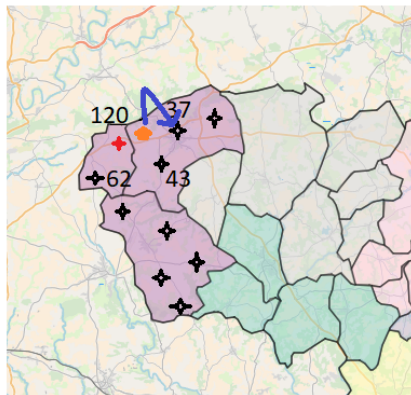
- Pour les RNP 2016 dans étage d'une RP 2016 :
grappe et secteur déduits automatiquement
 - Pour les RNP 2016 hors étage RP 2016 :
 - étape 1 : rattacher à un secteur
- méthode en cascade avec seuil maximal de 120 RNP par secteur



- ▲ RNP à rattacher à un secteur
- ✦ Barycentre de secteur
- ✦ Barycentre le plus proche de la RNP



Cas 1 : le secteur le plus proche est en-dessous de son seuil -> on lui rattache la RNP



Cas 2 : le secteur le plus proche a atteint son seuil -> on rattache la RNP au secteur le plus proche disponible

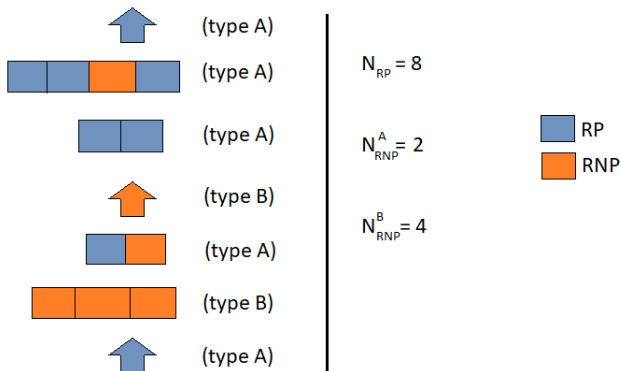
→ seuil à 400 pour cas atypiques (station de ski, station balnéaire...)

- Pour les RNP 2016 (suite) :
 - étape 2 : rattacher à la grappe la plus proche au sein du secteur
- Pour les logements nouveaux 2017 et 2018 : faibles volumes
 - étape 1 : rattacher au secteur le plus proche
 - étape 2 : rattacher à la grappe la plus proche au sein du secteur
- Obtention de « grappes-univers » 2018 :
 - en moyenne 25,11 logements dont 20,38 RP : satisfaisant
 - trop forte dispersion : le 99^e percentile vaut 67
 - beaucoup de grappes-univers sont trop grosses pour être collectées tel quel

Tirage stratifié d'étages

- Contraintes sur les RP :
 - Si moins de 24 RP dans la grappe-univers, les prendre toutes
 - Sinon en prendre 24
- Contraintes sur les RNP :
 - Si moins de 10 RNP dans la grappe-univers, les prendre toutes
 - Entre 11 et 40 RNP, en prendre 10
 - Entre 41 et 100, en prendre le quart
 - Plus de 100, en prendre 25
- Deux types d'étages :
 - type A : contient des RP et des RNP
 - type B : contient seulement des RNP

- Les RNP d'une grappe-univers se répartissent entre N_{RNP}^A et N_{RNP}^B



- Tirage stratifié d'étages : première phase

- Type A : $\pi_A = \min\left(1, \frac{24}{N_{RNP}}\right)$

→ apporte aussi des RNP, en moyenne $n_{RNP}^A = \pi_A \times N_{RNP}^A$

- Type B : $\pi_B = \max\left(\frac{n_{RNP} - n_{RNP}^A}{N_{RNP}^B}, \frac{1}{N_{RNP}^B}\right)$

- allocations évitant biais :

- $n_A = \max(\lfloor \pi_A N_A \rfloor, 1)$

- $n_B = \max(\lfloor \pi_B N_B \rfloor, 1)$

Problème : le total T de logements tirés peut dépasser 40

- Tirage stratifié d'étages : seconde phase

Si $T > 40$ et $n_E > 1$, une seconde phase enlève x_A étages de type A et x_B étages de type B

- Seconde phase :

- contrainte : $e := T - 40 = x_A \times m_A + x_B \times m_B$
(où m_A (resp. m_B) : nombre moyen de logement dans étages type A (resp. B))
- allocations théoriques de Neymann sous cette contrainte :

$$\begin{cases} x_A = n_A - \frac{n_A}{\sqrt{m_A}} \frac{n_A m_A + n_B m_B - e}{n_A \sqrt{m_A} + n_B \sqrt{m_B}} \\ x_B = n_B - \frac{n_B}{\sqrt{m_B}} \frac{n_A m_A + n_B m_B - e}{n_A \sqrt{m_A} + n_B \sqrt{m_B}} \end{cases}$$

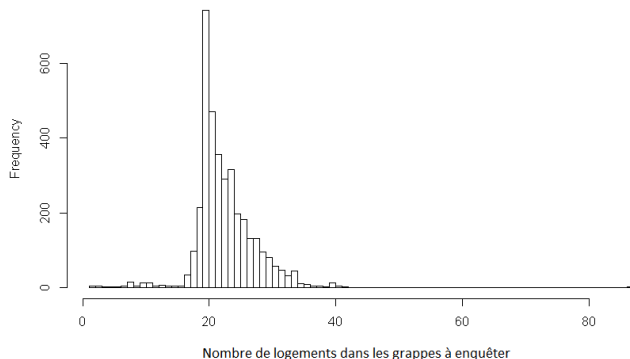
- allocations en pratique :

$$x'_A = \begin{cases} 0 & \text{si } n_A = 1 \\ \min(\lceil x_A \rceil, n_A - 1) & \text{sinon} \end{cases}$$

$$x'_B = \begin{cases} 0 & \text{si } n_B = 1 \\ \min(\lceil x_B \rceil, n_B - 1) & \text{sinon} \end{cases}$$

Résultats

- Histogramme typique des tailles des grappes à enquêter pour un trimestre donné :



- Pondérations : tous les logements d'un étage z ont même poids w .

Pour z situé dans une grappe g , laquelle appartient à un secteur s d'une UC U , en notant r le nombre de grappes (6 ou 7) du secteur, on a :

$$w = \underbrace{w_U}_{\text{obtenu par partage des poids}} \times \frac{1}{\pi_{s|U}} \times \underbrace{6}_{\text{affectation de } s \text{ à un des 6 trimestres d'entrée}} \times \frac{r}{6} \times \underbrace{6}_{\text{sélection de } g \text{ dans } s} \times \frac{1}{\pi_{z|g}}$$

- Bonne propriétés en pratique : la dispersion des poids est meilleure que pour l'ancienne EEC

Pour conclure

- Le nouvel échantillon de l'EEC provient d'un tirage à deux degrés
 - premier degré : tirage de secteurs proches de l'EM
 - second degré : tirage stratifié à deux phases d'étages dans des grappes-univers
- Les grappes à enquêter obtenues ont de bonnes propriétés :
 - respect des contraintes de nombre de RP et RNP
 - plus compactes que la précédentes EEC, notamment en zone rurale
 - une pondération moins dispersée qu'auparavant
- La Division Sondages a transmis les programmes de tirage à la Division Emploi

Merci pour votre attention !



DMCSI - DMS - Division Sondages