



ined
INSTITUT
NATIONAL
D'ÉTUDES
DÉMOGRA
PHIQUES

LES DOUBLES COMPTES DANS LES ENQUÊTES ANNUELLES DE RECENSEMENT EN FRANCE

Une estimation à partir de l'Echantillon
Démographique Permanent

Journées de méthodologie statistique de l'Insee, 12-14 juin 2018

Laurent Toulemon¹, Sébastien Durier², Benjamin Marteau¹

1: Institut national d'études démographiques

2: Institut national de la statistique et des études économiques

(version mise à jour après présentation)

Les doubles comptes dans les enquêtes annuelles de recensement en France

- Utiliser l'EDP pour estimer les doubles comptes au RP
 - Des doubles comptes observés, reconstitués et « invisibles »
 - Méthodes type « capture – recapture »
 - Valider les doubles comptes par un score d'appariement
- Caractéristiques des individus comptés deux fois
 - Age, situation familiale
- Impact sur l'estimation de la population
 - Nombre de ménages, nombre de personnes
- Conclusion
 - Recensement et autres données sociodémographiques
 - Post-stratification des enquêtes

Travaux dans le cadre du projet BigStat, https://big_stat.site.ined.fr



ined INSTITUT NATIONAL D'ÉTUDES DÉMOGRAPHIQUES

BIG_STAT

FR | EN

PROJET **SOURCES** **RECHERCHE** **LIENS**

🏠 Sources

Dans le projet Big_stat on utilise plusieurs sources de données: EDP: l'échantillon démographique permanent; ELFE: la cohorte d'enfants 2011; Recensement: les données individuelles du recensement; TCM: le fichier fusionné de la liste des habitants des ménages des enquêtes INSEE; DADS: les déclarations des employeurs et de l'Etat sur les salariés. La page «toutes les sources» indique comment accéder à des données ; les autres pages présentent les différentes sources sur lesquelles nous travaillons ou comptons travailler.

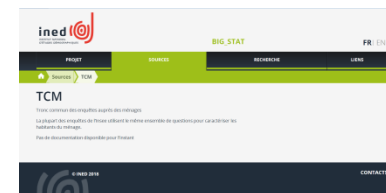
TOUTES LES SOURCES	SD-BOX	EDP
RECENSEMENT	ELFE	TCM

- Un projet financé par l'ANR, qui doit beaucoup à l'Insee
→ Relation entre utilisateurs (Ined, universitaires) et producteurs (Insee)

Le contenu du projet BigStat



- Recherche: population et situations familiales, jeunes adultes, enfants de parents séparés
 - Autres projets : orphelins, immigrés
 - Insee, Ined, universitaires
- Documentation des fichiers, formation des utilisateurs, retours vers les producteurs, collaborations
 - Les utilisateurs doivent prendre leur part du fardeau
- A l'Insee
 - Données Insee: EDP Recensement et EAR, Fideli, TCM
 - Mise à disposition, travaux en commun, retour sur les fichiers
- D'autres données: Elfe, Cnaf
- Certains sites sont actifs, d'autres pas encore



Utiliser l'EDP pour estimer les doubles comptes (1)

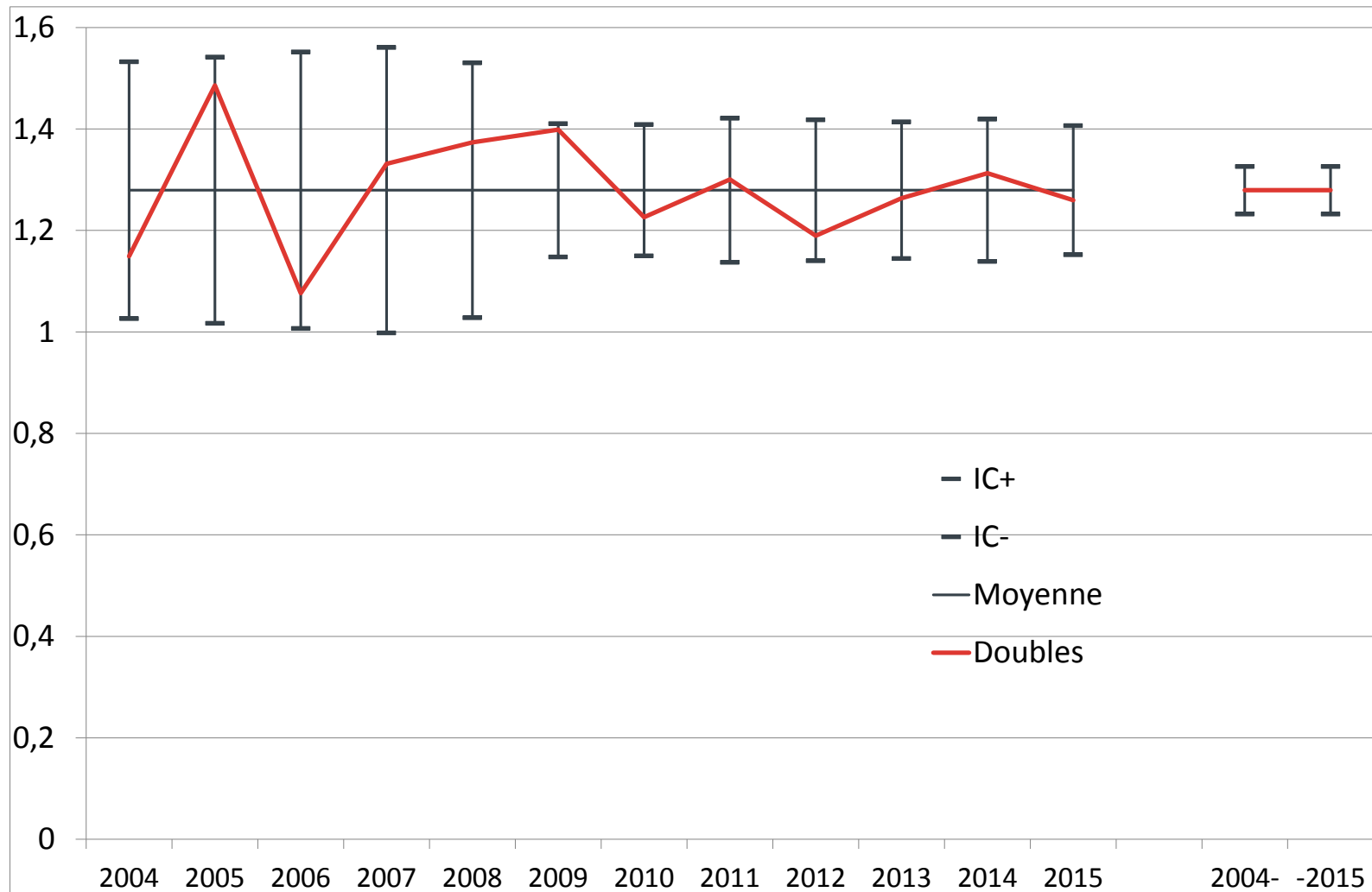
- Dans l'EDP: deux bulletins de recensement peuvent être associés à la même personne
- Probabilité d'être recensé
 - 20% dans les petites communes, 8% dans les grandes communes (environ, grandes adresses et logements neufs: 20%)
- Les probabilités des personnes EDP qui ont deux logements d'être recensées deux fois la même année sont très dispersées
 - Même petite commune : $20\% = 1/5$
 - La même grande commune, même adresse : $8\% = 2/25 = 1/12,5$
 - Deux petites communes différentes : $4\% = 1/5 * 1/5 = 1/25$
 - Une petite et une grande commune : $1,6\% = 1/5 * 2/25 = 1/62,5$
 - Deux grandes communes : $0,64\% = 2/25 * 2/25 = 1/156$
 - La même grande commune, deux adresses : $0,64\% = 1/156$
 - Grandes adresses, nouvelles adresses

Utiliser l'EDP pour estimer les doubles comptes (2)

- Trois types de doubles comptes des personnes recensées potentiellement dans deux logement
 - Recensés deux fois la même année (bulletins doubles)
 - Les habitants peuvent se rendre compte du double compte
 - Recensés deux fois en t et $t+1$ (reconstitués)
 - Question du lieu de résidence antérieure depuis 1 an
 - Présent dans un logement A en t , dans un logement B en $t+1$ et déclarant vivre dans ce même logement B depuis plus d'1 an
 - ➔ Alors également présent en B en t
 - Autres cas: recensé une fois en t
 - On fait l'hypothèse que la probabilité est la même que pour les personnes recensées en t et $t+1$
- Dans la plupart des cas (10 fois sur 11 environ) le double compte est invisible, mais bien réel

Doubles comptes en t , notés $(t;t)$

- 1,28% sur 2004-2015, 1,26% sur 2011-2015

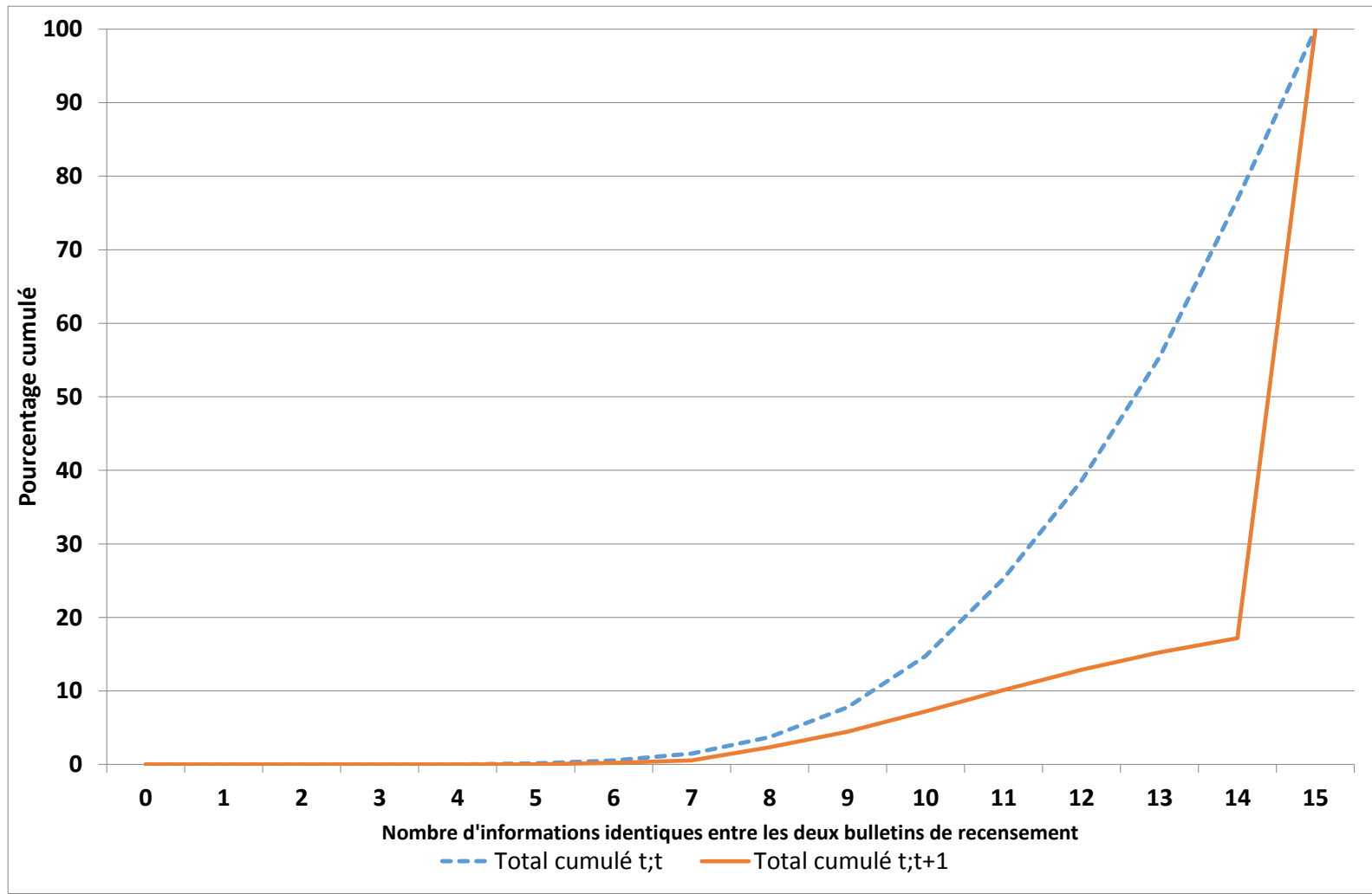


Doubles comptes en $t;t+1$

- Comparaison ($t;t+1$), à champ comparable
 - Champ observé : dans un logement en $t+1$ depuis plus d'1 an et dans un autre logement en t
 - sauf logements dans une même petite commune, même adresse dans une grande commune
 - Sauf déménagement (la moitié si un des deux logements)
 - 1,26% pour les années 2010 à 2014 en ($t;t$)
 - 0,76% dans le champ observé (homogène en $t;t$ et $t;t+1$)
 - **2,59% pour les années 2010 à 2014 en ($t;t+1$)**
 - + Après avoir tenu compte de la probabilité double pour deux années successives (A puis B ou B puis A)
 - On suppose que cette proportion s'applique à tous les cas où un seul logement est recensé une année donnée

Scores d'appariement entre deux bulletins: pas d'erreur sur la personne comptée deux fois

- Combinaison des deux bulletins de recensement et comparaison des informations renseignées : une meilleure concordance pour les individus recensés en $t;t+1$

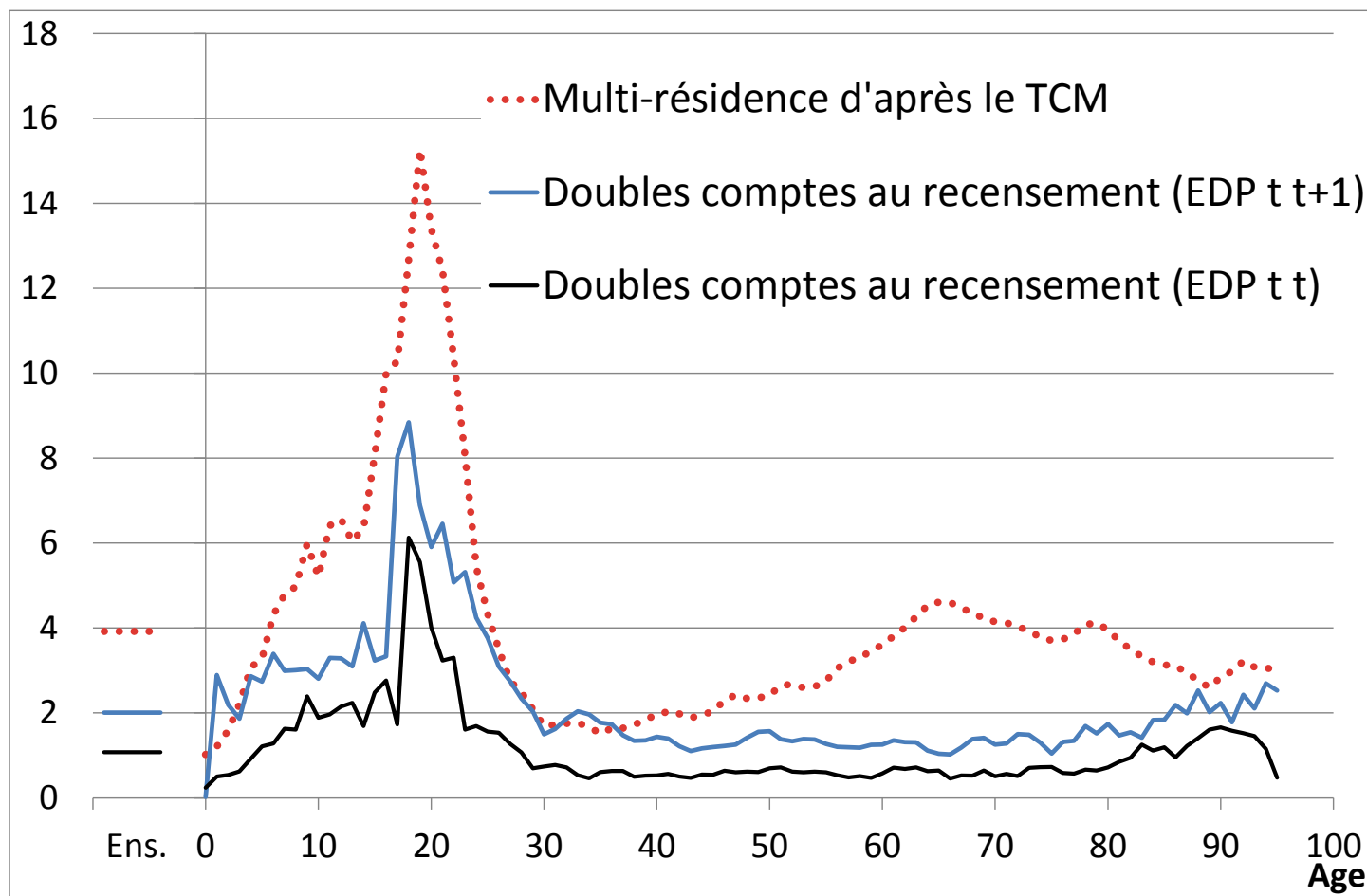


Combien de doubles comptes au recensement

- Une année donnée, dix personnes sur onze ayant deux logements sont recensées une seule fois (pas 12/13)
- Les doubles comptes sont deux fois plus probables quand les deux logements sont inclus dans le RP dans deux vagues successives
 - Précision de la variable « résidence un an avant » ?
 - **Estimation finale : 2,47% de doubles comptes**
 - Précision difficile à améliorer encore
- Profil des individus comptés deux fois
 - Age, situation familiale : profil général (2010-2014)
 - zoom sur les enfants des familles
- Correction pour les doubles comptes

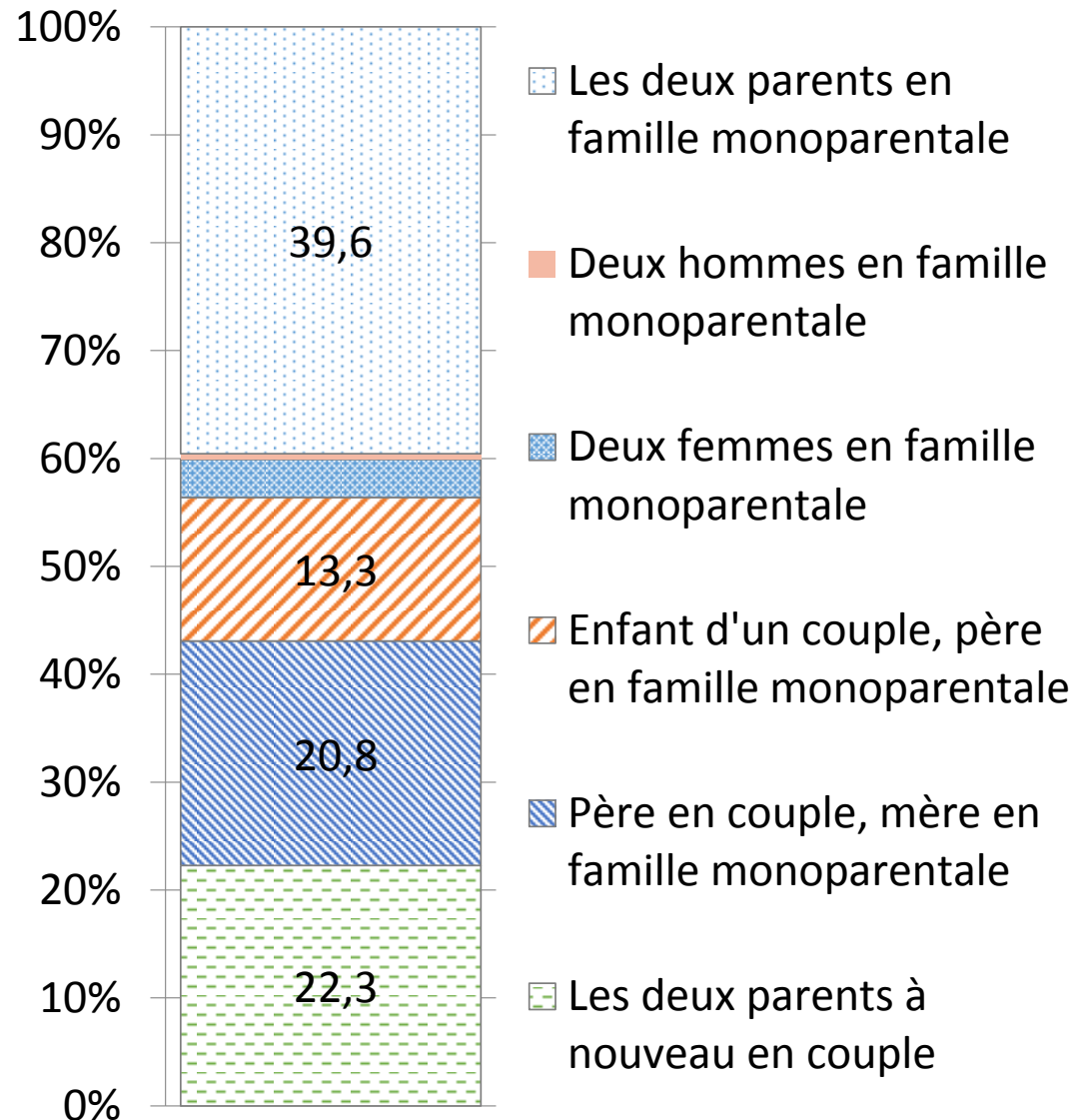
Profil par âge des individus en double compte

- Des proportions par âge assez proches de celles des personnes multi-résidentes dans le TCM



Enfants d'une famille dans les deux logements

- Enfants d'une famille dans les deux logements
 - 220 000 en t ; t
 - 330 000 en t ; $t+1$
- 2,44% des enfants mineurs
 - Les enfants ne remplissent pas eux-mêmes leurs bulletins
- Quel impact sur la répartition familiale?
 - On corrige des enfants présents deux fois



Vivre dans plusieurs configurations familiales



- *Insee Première N° 1647*

- Données de l'enquête Famille et logements
- Pas de correction pour les doubles comptes

Doubles résidence pour des enfants de parents séparés En milliers

Famille 1	Famille 2	Insee première 1647			Estimation EDP
		Rec+	Mono+	Moyenne	
Recomp	Recomp	290	-	145	128
Recomp	Mono	80	660	370	198
Mono	Mono	530	240	385	254
Total		900	900	900	580

- 900 000 enfants dans deux familles (mais 310 000 en trop)

- Une nouvelle estimation, en supprimant les doubles comptes

- 77,6% des enfants vivent avec leurs deux parents (et non pas 75,2%)

- Chiffres plus proches des TCM (80% corrigé des doubles comptes)

	En milliers		En %	
	IP1647	EDP	IP1647	EDP
Un parent	3 395	3 075	24,8	22,4
<i>Recomposée</i>	945	842	6,9	6,1
<i>Monoparentale</i>	2 450	2 233	17,9	16,3
Deux parents	10 305	10 625	75,2	77,6
<i>Recomposée</i>	531	547	3,9	4,0
<i>Simple</i>	9 774	10 078	71,3	73,6
Total	13 700	13 700	100	100

Quels types de corrections au recensement ?



- 2,5% d'habitants recensés à tort dans deux logements
- Ajuster pour les doubles comptes
 - Au recensement, la moitié des enfants en résidence partagée ont aussi été recensés dans leur autre logement
 - Au total 2,4% des enfants sont comptés dans deux familles
 - Cohérent avec le TCM: 4% des enfants mineurs ont deux résidences habituelles (après correction des pondérations)
- Autres questions
 - Autres bulletins remplis à tort (sans double compte)
 - Omissions
 - Impact sur la population totale ?
 - Liste des résidences principales à modifier également ?
 - Résidences principales, personnes par ménage

Conclusion

- Recensement et autres données sociodémographiques
 - Faire le lien avec d'autres fichiers
 - Partage des parts fiscales, multi-résidence dans les enquêtes
 - Omissions au recensement et présence fiscale
- Pondération des enquêtes avec TCM
 - Les habitants qui vivent aussi ailleurs
 - Partage des poids ou poids zéro une fois sur deux
- Collaborations en pratique
 - Diffuser les notes techniques sur les données
 - Travaux en commun ou échanges (méthodes et résultats)
 - Complémentarité entre producteurs et utilisateurs
- **Merci de votre attention**

Merci de votre attention

Contacts:

toulemon@ined.fr

sebastien.durier@insee.fr

benjamin.marteau@ined.fr