

Appariement du relevé structurel avec les données fiscales

Vers une statistique cantonale du revenu des ménages

Journées suisses de la statistique
12 novembre 2019

Matti Langel



Département des finances et des ressources humaines
Office cantonal de la statistique

Introduction et contexte

- Objectif : obtenir des indicateurs fiables concernant le revenu des **ménages** au niveau cantonal.
- Contexte : besoin et forte demande de résultats statistiques dans ce domaine.
- Seule la source fiscale permet d'approcher cette notion, mais imparfaitement car **contribuable ≠ ménage** :
 - 1 couple non marié avec 2 enfants majeurs = 4 contribuables distincts;
 - 1 couple marié avec 2 enfants mineurs = 1 contribuable.
- Corollaire : pas d'informations précises sur le taux de pauvreté, les inégalités, les conditions de vie à Genève.

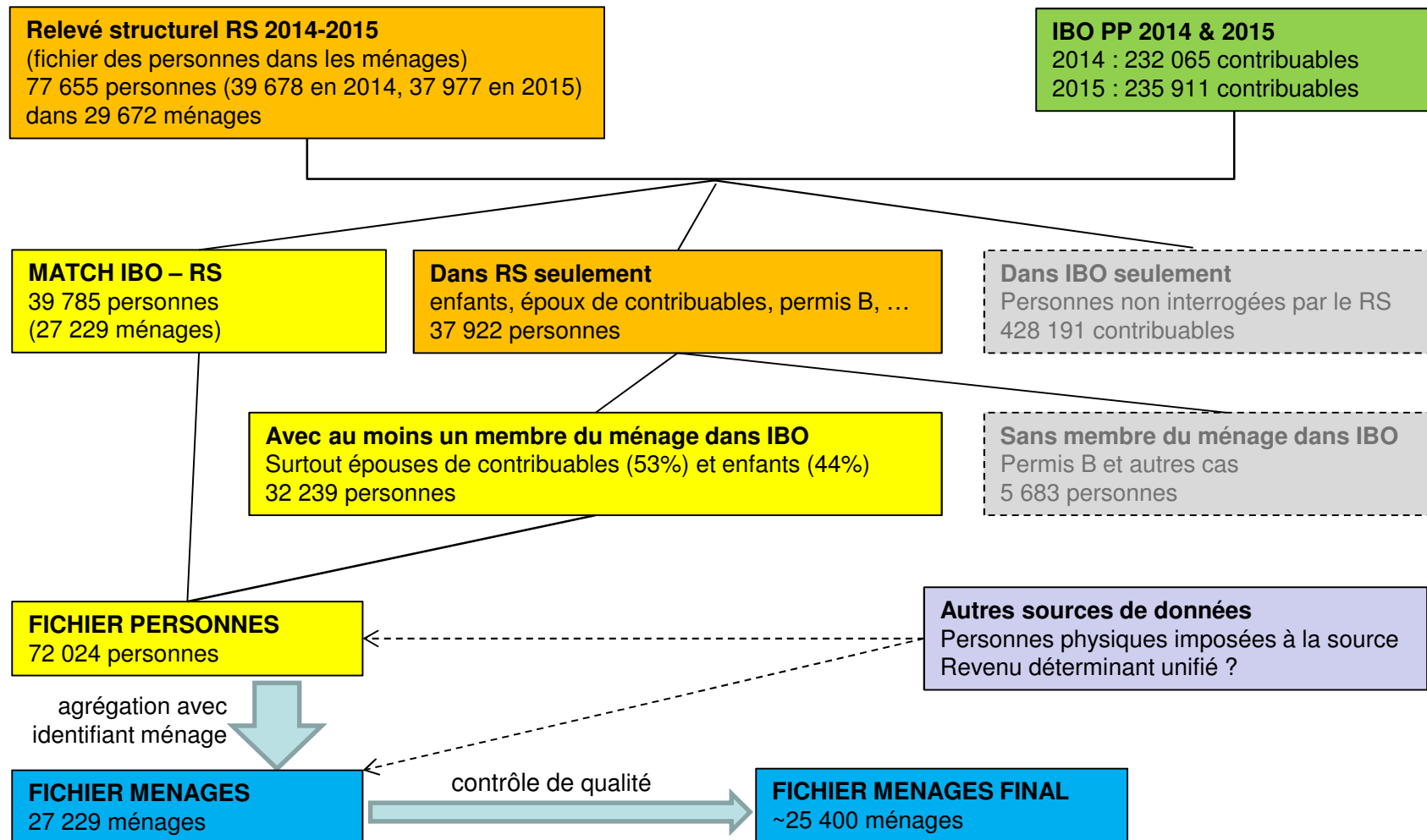
Solutions envisagées

1. Densification cantonale de l'enquête fédérale sur le revenu et les conditions de vie (SILC) :
 - petit échantillon (gros intervalles de confiance) malgré un coût élevé.
2. Appariement des données fiscales avec le registre des habitants :
 - qualité du registre pas encore suffisante à Genève en termes de ménages, peu de variables à disposition (données secondaires).
3. Appariement des données fiscales avec le relevé structurel (**solution retenue**) :
 - gros échantillon (densification cantonale), richesse de l'information (données primaires).

Appariement relevé structurel (RS) – données fiscales (IBO PP)

- Clé d'appariement : NAVS13 (doit être demandé à l'OFS).
- Base légale :
 - loi sur la statistique publique cantonale (LStat, du 24 janvier 2014);
 - loi sur la statistique fédérale (LSF du 9 octobre 1992).
- Premier appariement avec les données 2014 et 2015 :
 - données fiscales pers. physiques imposées au barème ordinaire (IBO PP);
 - relevé structurel : ménages et personnes dans les ménages (OFS);
 - autres : personnes physiques imposées à la source (IS PP).

Processus d'appariement



Quelques principes

- Relevé structurel : source de référence pour les variables sociodémographiques sur les personnes et les ménages.
- Données fiscales IBO : ne servent qu'à enrichir le RS avec les informations de revenu et fortune.
- Les revenus des personnes sont ensuite agrégés par ménage afin d'obtenir un revenu global du ménage.
- Données fiscales des personnes imposées à la source : lacunaires (ne contiennent que du salaire) donc utilisation minimale → compléter des revenus dans des ménages ayant au moins un contribuable IBO.

Qualité d'appariement (I)

- 51% des personnes dans le relevé structurel 2014-15 sont des contribuables IBO.
- Solde important (49%) pas étonnant : enfants mineurs, conjoints de contribuables couples mariés, contribuables imposés à la source (permis B essentiellement).
- Au final 93% des personnes dans le relevé structurel sont des contribuables IBO ou sont membre d'un ménage comportant au moins un contribuable IBO.
- **Risque (qualité)** : ne pas avoir les revenus de tous les membres du ménage (par ex. personne avec permis B) → sous-estimation du revenu du ménage.

Qualité d'appariement (II)

- Echantillon initial (RS) : ~ 29 600 ménages pour 2014-15.
- Ménages exclus de la base
 - *Si* aucun membre du ménage n'a été identifié comme contribuable imposé au barème ordinaire ayant résidé toute l'année dans le canton de Genève (~ 2 400 ménages exclus);
 - *ou si* au moins un adulte du ménage n'a pas été identifié dans les données fiscales (IBO ou personnes imposées à la source) (~ 1 600 ménages exclus);
 - *ou si* le revenu d'équivalence du ménage est inférieur au minimum vital pour une personne seule (1200.-/mois), (~ 200 ménages exclus).
- Echantillon final : ~ 25 400 ménages pour 2014-15.

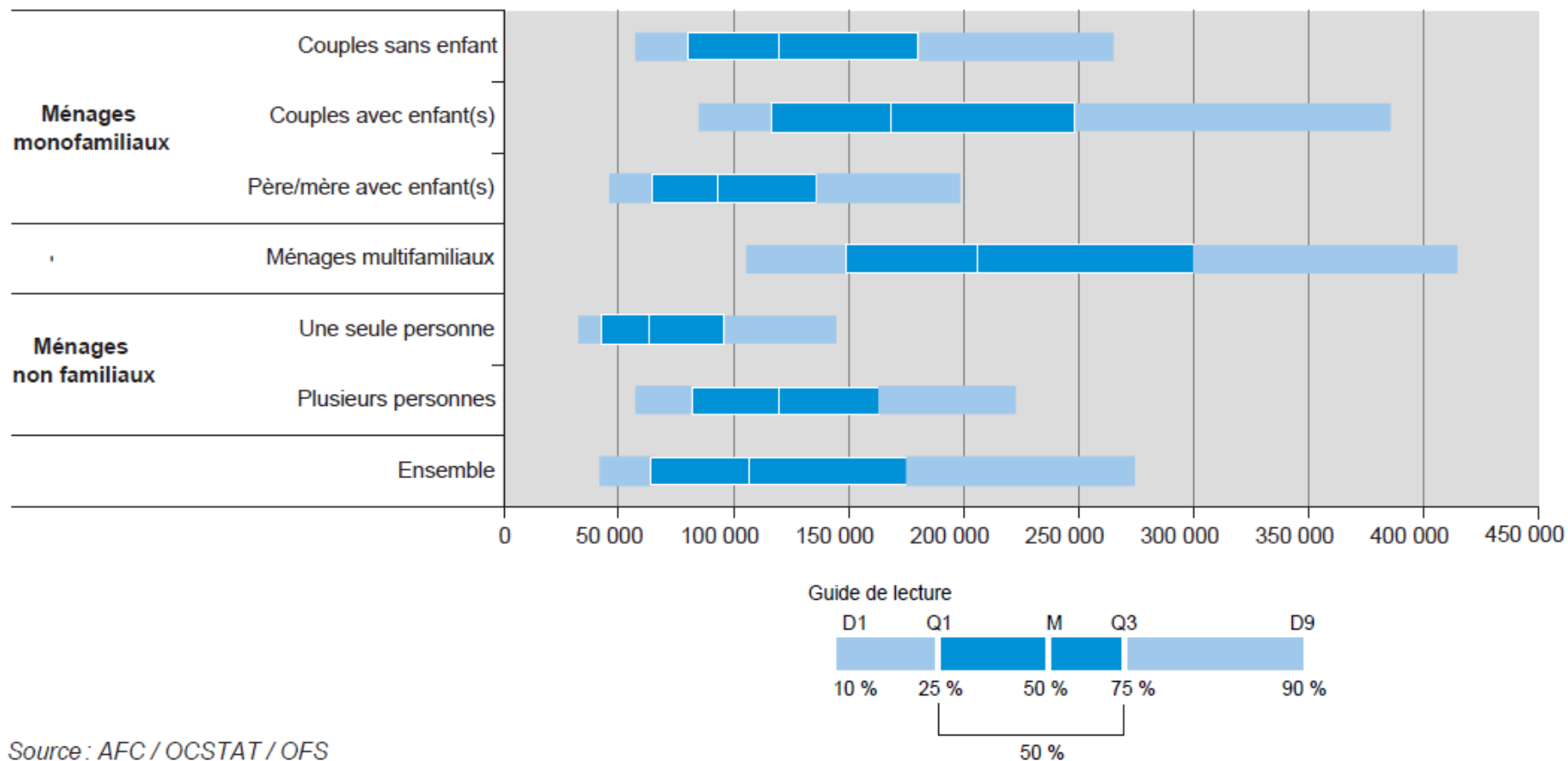
Pondérations

- Appariement d'un registre (IBO PP) avec un échantillon (RS).
- Plan de sondage du RS : chaque ménage a un poids spécifique.
- L'exclusion d'environ 15 % des ménages n'est pas forcément neutre en termes de structure.
- Après analyse de plusieurs variables de structure des ménages, pas de repondération opérée (pour l'instant), sauf pour une valeur extrême.

Possibilités d'analyse et nouveaux indicateurs

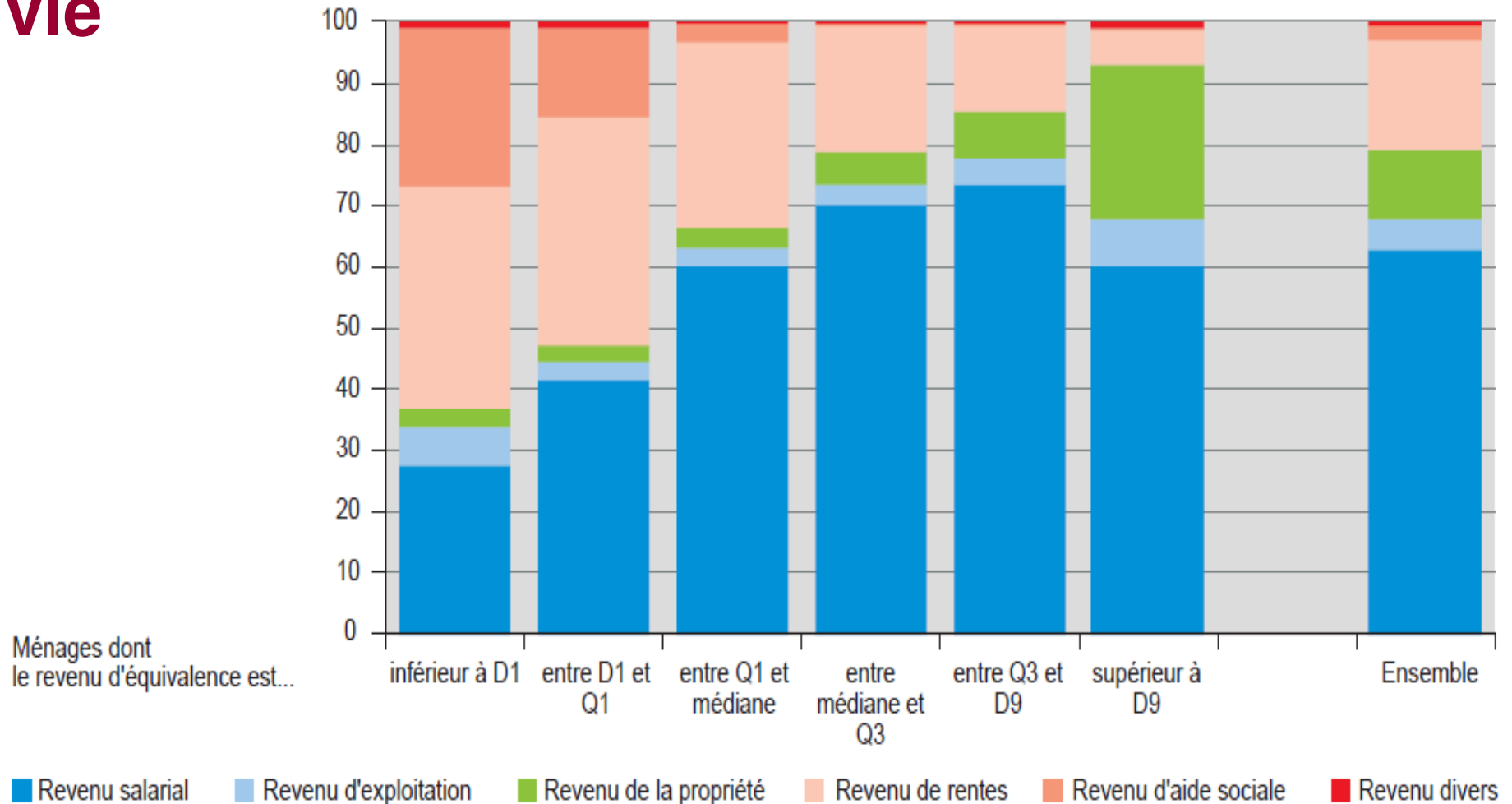
- Distribution de revenus, quantiles.
- Calcul et utilisation d'un revenu d'équivalence.
- Décomposition du revenu (types de revenus).
- Indicateurs d'inégalités et de pauvreté.
- Beaucoup d'informations sur les ménages grâce au RS :
 - logements;
 - types de ménages, de familles, etc.;
 - nationalités, langues;
 - statut sur le marché du travail;
 - niveau de formation;
 - etc.

Revenu brut des ménages en 2014-15 dans le canton de Genève, par type de ménage



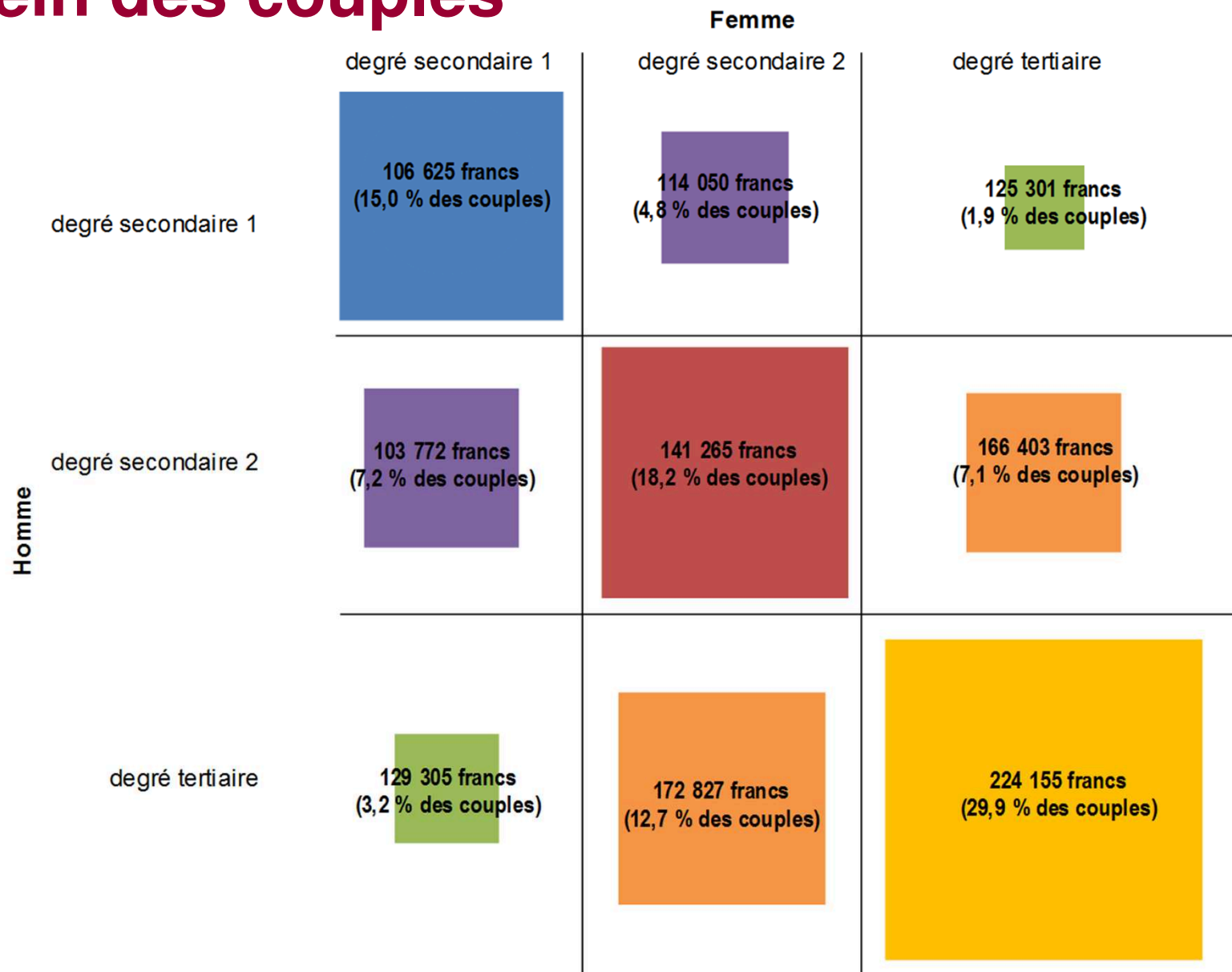
Source : AFC / OCSTAT / OFS

Composition du revenu selon le niveau de vie



Source: AFC / OCSTAT / OFS

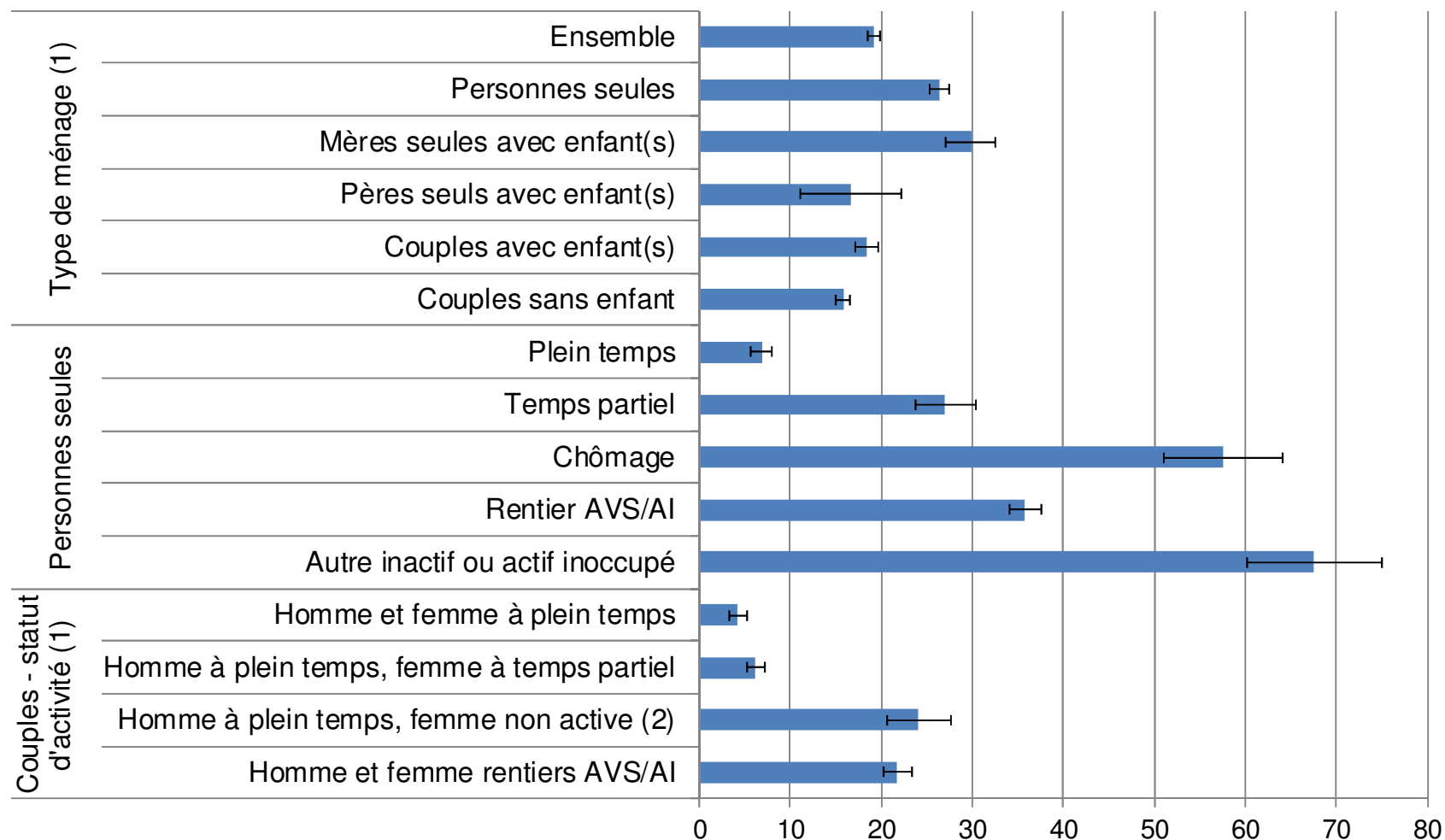
Revenu selon le niveau de formation au sein des couples



Taux de risque de pauvreté (à 60% de la médiane)

En %

Canton de Genève



(1) Configuration les plus fréquentes.

(2) Hors rentière AVS/AI et chômeuse.

Source : AFC / OCSTAT / OFS

Distribution du revenu et mesures d'inégalités (revenu brut d'équivalence)

	Genève 2014-15	Suisse 2015
Quantiles principaux		
1er décile (D1)	35 057	35 976
1er quartile (Q1)	48 353	49 776
Médiane (M)	72 657	67 380
3e quartile (Q3)	106 323	91 284
9e décile (D9)	157 535	125 340
Mesures d'inégalités		
Rapport interquartile (IQR)	2,2	1,8
Rapport interdécile (IDR)	4,5	3,5
Indice de Gini	0,38	0,29
Rapport S80/S20	6,4	4,3

Source : OCSTAT - Statistique du revenu des ménages / OFS - EBM

Valeur extrême (I)

- Dans les données RS figure le ménage avec le revenu le plus élevé du canton.
- Revenu ~ 50-100 fois supérieur aux autres revenus les plus importants des contribuables genevois.
- Son poids dans le RS fait qu'il « représente » environ 5 ménages.
- Problème : effet très important sur l'estimation de tous les indicateurs non robustes (indice de Gini, S80/S20).

Valeur extrême (II)

- Solution : repondération de cette observation → poids=1
- Justification : il faut conserver ce ménage car il existe. Grâce aux informations auxiliaires (données fiscales de l'ensemble des contribuables), on sait qu'il est unique en termes de niveaux de revenus.

	traitement de la valeur extrême		
	Repondération (poids=1)	Pas de repondération	Suppression
Gini	0,38	0,42	0,37
S80/S20	6,4	7,3	6,2

Limites

- Il manque des groupes entiers de population :
 - les étrangers non permis C sont hors champ IBO ; données IS trop lacunaires pour compléter;
 - les fonctionnaires internationaux, les diplomates et les membres de leur famille sont hors champ RS.
- Pas encore de possibilité de calculer un revenu disponible (soustraction des dépenses obligatoires), donc comparabilité moindre (enquête SILC, résultats d'autres cantons, etc.).
- Données d'enquête : fluctuations d'échantillonnage peuvent rendre difficile l'analyse des évolutions.

Développements futurs (I)

- Mise à jour des indicateurs de manière régulière, avec données « poolées » sur 3 ans.
- Intégration des données du RDU cantonal :
 - Amélioration de la qualité des revenus d'aide sociale ;
 - Prise en compte plus large des personnes imposées à la source.
- Estimation d'un revenu disponible pour améliorer la comparabilité (inter cantonale notamment)
- Adaptation de la pondération en exploitant l'information auxiliaire à disposition (calage)

Développements futurs (II)

- Calcul d'un taux de pauvreté (seuil absolu en adaptant les normes CSIAS).
- Calcul d'intervalles de confiance pour l'ensemble des indicateurs d'inégalités et pauvreté.
- Exploitation de tout le potentiel du RS en termes d'information disponible sur les ménages (publications sur des thèmes spécifiques).
- Projet évolutif : modèle de données flexible plutôt que fichier statique.

Merci de votre attention