

**Commission économique pour l'Europe**

Conférence des statisticiens européens

**Groupe d'experts des recensements de la population et des habitations****Vingt-cinquième réunion**

Genève, 20-22 septembre 2023

Point 2 de l'ordre du jour provisoire

**Enseignements à retenir du cycle des recensements de 2020****Recensement estonien de 2021 : utilisation de graphes pour déterminer les ménages et les logements à partir des registres****Note de Statistics Estonia\*<sup>1</sup>***Résumé*

Le recensement de 2021 a été le premier en Estonie à produire toutes les caractéristiques obligatoirement requises dans le cadre de l'UE à partir de données administratives. L'un des principaux défis à relever était la faible précision (à peine 80 %) des données relatives au lieu de résidence inscrites dans le registre de la population et les conséquences de cet état de fait sur les statistiques relatives aux ménages.

Dans un recensement fondé sur les registres, le ménage est défini comme un ensemble de personnes vivant dans le même logement. Lorsqu'on utilise le lieu de résidence indiqué dans le registre de la population pour déterminer les ménages et les familles, les statistiques obtenues surestiment le nombre de parents isolés et sous-estiment le nombre de couples.

Pour améliorer les statistiques sur les ménages et les familles, nous avons mis au point une méthode qui s'appuie sur des graphes et utilise des données provenant de sources administratives. Nous considérons que les personnes et les adresses constituent les sommets d'un graphe. Les liens entre deux personnes (comme le mariage ou la parenté) ou entre une personne et un lieu (comme la propriété d'un bien immobilier) constituent les arêtes du graphe. Un ménage est considéré comme un graphe secondaire contenant les membres du ménage et leur logement. Déterminer les ménages et leurs logements revient donc à trouver

\* Document établi par Helle Visk.

*Note* : Les appellations employées dans le présent document n'impliquent de la part du Secrétariat de l'Organisation des Nations Unies aucune prise de position quant au statut juridique des pays, territoires, villes ou zones, ou de leurs autorités, ni quant au tracé de leurs frontières ou limites.

<sup>1</sup> Le présent document a été soumis avec retard en raison de sa transmission tardive par Statistics Estonia.



des graphes secondaires densément connectés ou, en d'autres termes, à détecter des communautés.

Pour trouver des liens entre les personnes ou entre les personnes et les lieux, nous avons utilisé les données de 17 registres. Chaque arête du graphe a été affectée d'un coefficient de pondération décrivant la probabilité que des personnes vivent ensemble ou qu'une personne vive à une adresse donnée. Les modèles de probabilité ont été établis en fonction des données sur les ménages provenant d'enquêtes existantes.

Le nouveau cadre a été utilisé pour comptabiliser les ménages et les lieux de résidence dans le recensement.

## I. Introduction

1. Le recensement de 2021 a été le premier recensement effectué en Estonie dans lequel toutes les variables obligatoirement requises dans le cadre de l'UE ont été obtenues à partir de données administratives. Le système de registres estonien se prête parfaitement à l'organisation d'un recensement fondé sur les registres. En effet, ceux-ci couvrent une grande diversité de sujets liés au recensement. En outre, il est facile de relier les sources, car nous disposons d'identifiants uniques pour les personnes, les adresses et les entreprises.

2. Bien que la qualité des données des registres soit généralement élevée, il existe quelques exceptions. Par exemple, les données relatives au lieu de résidence qui figurent dans le registre de la population ne sont exactes que pour environ 80 % des personnes (Gortfelder & Puur, 2021). Les raisons pour lesquelles certaines personnes ne mettent pas à jour les informations les concernant dans le registre sont, entre autres, que ces personnes considèrent l'enregistrement comme inutile, qu'elles utilisent les services et les avantages d'une municipalité donnée ou qu'elles considèrent leur domicile actuel comme temporaire (Gortfelder & Puur, 2021 ; Äär, 2017).

3. Dans un recensement fondé sur les registres, un ménage est composé de personnes vivant à la même adresse, qu'elles aient ou non un budget commun. La famille est définie au sens étroit, comme un noyau familial. Il s'agit soit d'un couple en concubinage ou marié avec ou sans enfants, soit d'un parent isolé avec un ou plusieurs enfants. La famille se compose de personnes vivant au sein du même ménage. Le lieu de résidence étant la base de la répartition de la population en ménages, son inexactitude a une incidence sur les statistiques relatives aux ménages et aux familles.

4. Dans le recensement pilote réalisé par Statistics Estonia en 2016, les ménages ont été déterminés à partir des données relatives au lieu de résidence inscrites dans le registre de la population. Les statistiques sur les ménages et les familles différaient considérablement de celles qui avaient été établies lors du recensement de 2011. Par exemple, le nombre de parents isolés était supérieur de 67 % par rapport au recensement de 2011, le nombre de partenaires était inférieur de 26 %.

5. La surestimation du nombre de parents isolés est caractéristique d'une situation dans laquelle les membres de la famille sont enregistrés à des adresses différentes. Considérons par exemple une famille de quatre personnes : la mère, le père, la fille et le fils. Si le père et la fille s'inscrivent à une adresse différente, ils apparaîtront comme deux familles monoparentales (mère-fils, père-fille) dans le registre de la population.

6. Les statistiques sur les ménages et les familles dérivées du registre de la population étant fortement biaisées, il n'était pas possible d'utiliser les données relatives au lieu de résidence pour le recensement proprement dit. Pour obtenir de meilleures statistiques, nous devons trouver une méthode pour réunifier les familles qui apparaissaient dispersées dans les registres.

## II. Méthodes

### A. Ménages ou logements

7. Reconstituer les familles ne se limite pas à trouver des liens entre les membres éventuels d'une même famille. Reprenons l'exemple du paragraphe 5. La réunification de cette famille implique que le père et la fille « réintègrent » le domicile de la mère et du fils. Il convient de noter que cela suppose d'assigner au père et à la fille un logement différent de celui qui est constaté dans le registre de la population. En outre, la composition du ménage serait différente de celle qu'indique le registre pour l'ensemble des membres de la famille. Par conséquent, la réunification des familles implique une modification de la composition des ménages et de la localisation des logements.

8. Les informations permettant de déterminer les ménages et leurs logements peuvent être recueillies à partir de sources administratives. Par exemple, les données sur la parenté et les mariages permettent de savoir quels sont les membres d'une même famille. Les données sur les contrats de propriété et d'électricité peuvent être utiles pour localiser les domiciles des uns et des autres.

9. Il est difficile de savoir s'il est préférable de commencer par déterminer les ménages puis de leur assigner à chacun un logement, ou de localiser d'abord le domicile de chaque résident, puis de considérer les ménages constitués de personnes vivant à la même adresse. Si l'on commence par les ménages, on passe à côté des familles qui n'ont pas de liens directs, comme les concubins qui sont copropriétaires d'un appartement. En revanche, si l'on relie des personnes à des adresses en ignorant les données qui relient ces personnes (par exemple, le mariage), il est peu probable que les familles qui en résultent correspondent beaucoup mieux à la réalité que dans le registre de la population.

### B. Réseau de personnes et de lieux

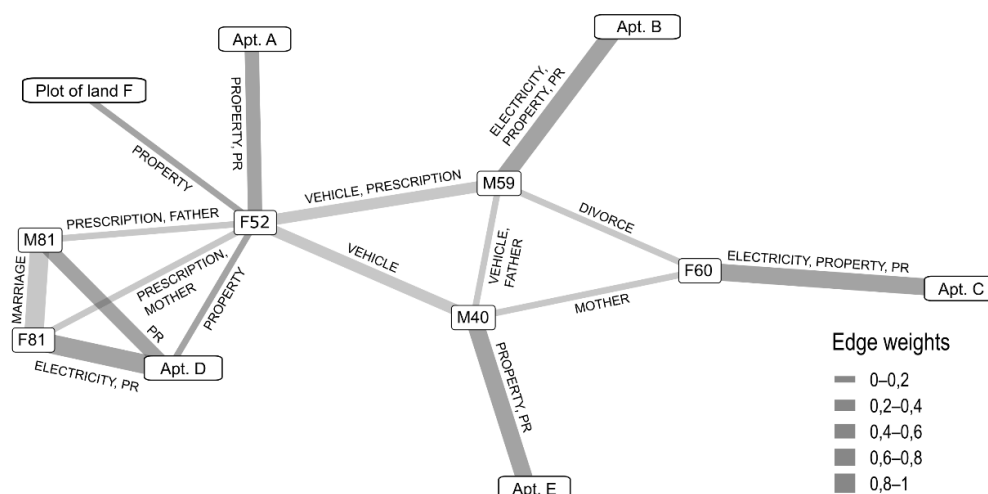
10. Notre idée est de simultanément constituer des ménages et assigner des logements. Pour cela, nous considérons les personnes et les logements comme les sommets d'un graphe. Les arêtes de ce graphe sont figurées par les éléments suivants :

- Les liens entre les personnes (tels que le mariage, la parenté, l'achat de médicaments sur ordonnance pour quelqu'un, le partage d'une voiture) ; ou
- Les liens entre les personnes et les lieux (lieu de résidence, propriété, contrat d'électricité, par exemple).

Un exemple de ce graphe est présenté à la figure 1.

Figure 1

#### Extrait du graphe des personnes et des lieux



*Note : Les sommets sont des personnes (la vignette indique le sexe et l'âge) et des lieux (des appartements et une parcelle de terrain). Les arêtes relient des personnes entre elles (gris clair) ou des personnes à des lieux (gris foncé). Le coefficient de pondération (Edge weight) des arêtes indique la probabilité que les personnes considérées vivent au sein du même ménage ou qu'elles habitent à la même adresse. Les libellés des arêtes indiquent le type de relation. ELECTRICITY : la personne a un contrat d'électricité à l'adresse indiquée, VEHICLE : les personnes sont liées au même véhicule en tant que propriétaires ou utilisateurs, PRESCRIPTION : la personne a acheté un médicament sur ordonnance pour l'autre personne, PR : lieu de résidence dans le registre de la population. Cette figure a été publiée précédemment (Tiit et al., 2021).*

11. Nous nous intéressons à la division du graphe de l'ensemble des personnes et des logements en graphes secondaires contenant chacun les membres d'un ménage et leur logement. Il est naturel de supposer que les membres d'un même ménage sont fortement liés entre eux et ont un lien étroit avec leur logement.

12. Dans la théorie des graphes, une communauté est définie comme un groupe de sommets qui sont assez densément connectés entre eux mais peu connectés à d'autres groupes denses du réseau (Porter et al., 2009). Nous en concluons que déterminer les ménages et les logements revient à déterminer des communautés. La détection des communautés est une tâche courante dans l'analyse des réseaux et de nombreux algorithmes sont disponibles dans divers langages de programmation.

13. Les liens entre les personnes, ou entre les personnes et les logements, sont plus ou moins forts. Par exemple, presque tous les enfants mineurs vivent avec leur(s) parent(s), mais cela est moins fréquent lorsque l'enfant devient adulte. Il est donc logique d'attribuer des coefficients de pondération aux arêtes.

## C. Déroulement des tâches

14. La première étape consiste à collecter les données. Nous appelons les différents types de connexion des signes de personne à personne ou de personne à lieu. Les données sur les signes ont été obtenues à partir de 17 registres (tableaux 1 et 2). En outre, les données provenant de grandes enquêtes annuelles sur les ménages (Enquête sociale estonienne/Statistiques de l'UE sur le revenu et les conditions de vie (EU-SILC) de 2021, Enquête sur la population active en Estonie de 2021 – ont été utilisées comme données d'apprentissage pour modéliser la relation entre les signes et les situations observées dans la vie réelle.

Tableau 1  
Signes de personne à personne issus des registres

Registre	Signes de personne à personne
Fichier électronique	Les personnes sont du même côté dans un litige portant sur une pension alimentaire  Les personnes se trouvent dans des camps opposés dans un litige portant sur une pension alimentaire
Système d'information sur l'assurance maladie	Une personne s'est occupée d'une autre personne au cours de l'année précédant le recensement
Registre de la circulation	Les personnes sont liées au même véhicule (par exemple, l'utilisateur et le propriétaire de la voiture)
Registre des contribuables	Les personnes ont déposé une demande conjointe de prêt hypothécaire  L'un des conjoints a transféré des revenus non imposables à l'autre conjoint

<i>Registre</i>	<i>Signes de personne à personne</i>
	Utilisation de l'avantage fiscal pour deux enfants et plus (lien entre l'enfant et la personne qui soumet la déclaration)
	Utilisation de l'avantage fiscal pour couvrir les frais d'études d'un tiers
Centre estonien de prescription médicale	La personne a acheté les médicaments sur ordonnance prescrits à une autre personne
Registre de la population	Les personnes sont mariées
	Les personnes sont divorcées
	Un adulte sert de tuteur à un autre adulte
	La personne est la mère de l'autre personne
	La personne est le père de l'autre personne
	La personne a un droit de garde complet sur l'enfant
	La personne a un droit de garde limité sur l'enfant
	L'enfant est séparé de son parent
Registre des prestations et services sociaux	Les personnes ont bénéficié d'une allocation de subsistance au sein du même ménage
Système d'information sur la sécurité sociale	La personne perçoit des allocations familiales pour un enfant
	La personne perçoit une allocation parentale pour un enfant
	Un adulte bénéficie d'un congé supplémentaire pour s'occuper d'un adulte handicapé

Tableau 2  
**Signes de personne à lieu issus des registres**

<i>Registre</i>	<i>Signes de personne à lieu</i>
<b>Logements potentiels</b>	
Elering (exploitant du réseau électrique)	La personne a un contrat d'électricité à l'adresse considérée
Registre des personnes inscrites comme chômeurs ou demandeurs d'emploi et de la prestation de services d'aide à l'emploi	Lieu de résidence de la personne Adresse postale de la personne
Registre d'écrou	Lieu de résidence des probationnaires
Registre foncier	Biens immobiliers appartenant à la personne
Registre de la population	Lieu de résidence déclaré de la personne Adresse additionnelle de la personne

<i>Registre</i>	<i>Signes de personne à lieu</i>	
	Précédents lieux de résidence de la personne	
	Lieu de séjour de la personne (exemple : dortoir)	
Recensement de la population et des habitations de 2011	Adresses de la personne et de sa mère	
Registre des prestations et services sociaux	Lieu de résidence de la personne	
Registre des contribuables	Biens immobiliers achetés avec un prêt personnel au logement	
<b>Niveau de la municipalité</b>	Système estonien d'information sur l'éducation	École maternelle de l'enfant
		Étudiant à l'université ou dans un établissement d'enseignement professionnel
		École de l'élève dans l'enseignement général
		Lieu de travail de l'enseignant
Système d'information sur l'assurance maladie		Institution de soins dentaires fréquentée par la personne
		Institution médicale fréquentée par la personne
		Médecin traitant de la personne
Base de données des documents d'identité	Lieu de réception d'un document d'identité	
Registre du fonds de retraite obligatoire	Adresse de la personne qui a adhéré au fonds de retraite	
Centre estonien de prescription médicale	Pharmacie dans laquelle la personne a acheté des médicaments	
Registre de l'emploi	Lieu de travail de la personne	

15. La pondération des arêtes a été modélisée comme suit :

a) Probabilité que les personnes vivent au sein du même ménage (un modèle de régression logistique a été appliqué) ; ou

b) Probabilité que la personne vive dans un logement (forêt aléatoire). Ce dernier modèle inclut également des données au niveau de la municipalité (comme le fait d'avoir un médecin généraliste dans une municipalité donnée), ainsi que les distances par rapport au jardin d'enfants, à l'école et au lieu de travail.

Les modèles ont été adaptés aux données de l'enquête, puis appliqués à l'ensemble de la population.

16. Le graphe comprenait au total 5,2 millions de sommets et 7,8 millions d'arêtes. La détection des communautés a été appliquée en deux phases :

a) Tout d'abord, la méthode de Louvain (Blondel et al., 2008) a été employée pour fractionner le graphe initial en graphes secondaires ne comportant pas plus de 5 000 sommets ;

b) Dans chacun de ces graphes secondaires, l'algorithme Infomap (Rosvall & Bergstrom, 2008) a été appliqué de manière récursive jusqu'à ce que les communautés soient suffisamment réduites ou qu'il n'y ait plus d'augmentation importante de la modularité.

17. Les communautés ainsi obtenues étaient similaires à des ménages et les statistiques sur les familles se sont trouvées améliorées par rapport à celles qui étaient dérivées du registre de la population (tableau 3, fig. 2). Toutefois, le nombre de ménages multifamiliaux et de familles comprenant des enfants adultes était surestimé, tandis que le nombre de ménages composés d'une seule personne était sous-estimé. En outre, quelques enfants se retrouvaient au sein de ménages sans adultes.

Tableau 3  
**Situation familiale des personnes dans les données d'apprentissage**

<i>Situation familiale</i>	<i>Données pilotes</i>	<i>Registre de la population</i>	<i>Grappes</i>
Partenaires, mariés	34,6	27,6	34,6
Partenaires, en concubinage	16,3	10,4	14,5
Parents isolés	4,2	9,6	4,8
Enfant, non issu d'un parent isolé	24,2	19,6	27,3
Enfant, issu d'un parent isolé	5,6	13,5	5,9
N'est pas au sein d'un noyau familial	15,1	19,3	12,9

*Note : Les données d'apprentissage sont issues des statistiques EU-SILC de 2021 et de l'enquête sur la population active de 2021.*

*Le tableau inclut les données de 32 802 personnes qui figuraient dans toutes les sources (ce qui exclut les non-résidents, membres de ménages institutionnels à la date de recensement du 31 décembre 2021 à 00 h 00, les personnes nées après cette date et celles qui sont décédées avant). Les données ne sont pas pondérées.*

18. Lors du traitement a posteriori, les enfants vivant seuls ont été ajoutés aux ménages composés de leurs parents ou d'un autre adulte apparenté. À l'aide des données de l'enquête, une heuristique a été mise au point pour fractionner certaines des communautés les moins connectées (par exemple, s'il y avait plusieurs familles dans un ménage, nous avons choisi de séparer la famille la moins connectée pour l'inclure dans un autre ménage).

19. Au cours de la phase suivante, chaque ménage a été affecté à un logement. Cette tâche est très simple s'il n'y a qu'un seul logement dans la communauté. Cependant, certaines communautés pouvaient être rattachées à plusieurs logements, d'autres à aucun. En outre, à l'issue du traitement a posteriori décrit au paragraphe 18, certaines communautés comportaient plusieurs ménages entrant en concurrence pour occuper le ou les logements disponibles. En général, chaque communauté comprenait  $m \geq 0$  logement et  $n \geq 1$  ménage.

a) Dans chaque communauté, le degré de connexion entre les ménages et les logements a été calculé à partir des pondérations entre les personnes et les logements. Nous avons privilégié les combinaisons ménage-logement les plus fortement connectées. En cas d'égalité, nous avons donné la priorité aux logements présentant une consommation d'électricité plus élevée et aux logements de plus grande taille. Au cours de cette étape, 96 % des ménages se sont vu assigner un logement :

i) Pour le reste des ménages, la municipalité la plus probable a été sélectionnée. En outre, nous avons calculé un point d'ancrage pour chaque ménage dans la municipalité choisie, sur la base des coordonnées géographiques des lieux avec lesquels les membres du ménage étaient en relation ;

ii) Les habitations candidates à ce stade ont été sélectionnées parmi les logements laissés inoccupés. Nous avons pris en considération les lieux avec lesquels les membres du ménage étaient en relation, soit directement, soit par l'intermédiaire d'autres personnes. Un exemple de logement potentiel pourrait être l'appartement de

la mère d'un membre du ménage. Pour chaque ménage, nous n'avons pris en compte que les logements situés dans la municipalité sélectionnée ;

iii) Il y avait une relation multivoque entre les ménages et les logements. Certains ménages avaient plusieurs candidats raisonnables pour le logement. À l'inverse, certains logements étaient des candidats pour différents ménages. Dans de tels cas, il existe de nombreuses façons de faire correspondre les ménages et les logements. Nous avons considéré les ménages et les logements comme un graphe bipartite et avons opté pour un appariement stable, forme d'appariement dans laquelle il n'y a pas de combinaison ménage-logement qui aurait la préférence par rapport à ses « compagnons » (Gale & Shapley, 1962) ;

iv) Une condition préalable au calcul de l'appariement stable est une sorte de classement : quels ménages préfèrent quels logements et vice versa. Nous avons posé comme principe que les ménages préféreraient les logements qui étaient 1) proches de leur point d'ancrage et 2) plus grands. Pour les logements, la préférence allait 1) aux ménages les plus nombreux et 2) aux ménages les plus proches ;

À l'issue de cette étape, 99,4 % des ménages étaient rattachés à un logement ;

b) Les autres ménages se sont vu attribuer un logement inoccupé choisi au hasard à proximité de leur point d'ancrage.

20. Les personnes vivant au sein des ménages institutionnels et les sans-abri ont été traités séparément. Les municipalités ont fourni des listes de sans-abri ; les registres ont permis de connaître les membres des ménages institutionnels.

### III. Résultats

21. Les statistiques sur les familles et les ménages établies dans le cadre du recensement ont été calculées à l'aide de la nouvelle méthode. Celle-ci a par ailleurs remplacé le registre de la population comme base de la ventilation géographique des statistiques démographiques annuelles à partir du 1<sup>er</sup> janvier 2022.

22. Malgré l'utilisation de différentes sources relatives aux logements, 3 personnes sur 4 ont conservé leur lieu de résidence tel qu'il était indiqué dans le registre de la population. Comme certaines personnes se sont vu assigner des logements différents, la population des municipalités a changé par rapport aux chiffres du registre de la population. Sur 79 municipalités, 35 sont restées à peu près aussi peuplées que selon le registre ( $\pm 2\%$ ), 23 ont perdu au moins 2 % de leur population et 21 ont enregistré un gain de plus de 2 %. Les pertes les plus importantes ont été observées dans les petites îles (Ruhnu -29 %, Vormsi -23 %, Kihnu -19 %, etc.) et dans d'autres régions populaires en tant que lieux de résidence d'été (Alutaguse -8 %, Narva-Jõesuu -6 %). Les villes russophones du nord de l'Estonie (Loksa +7 %, Maardu +6 %, Kohtla-Järve +4 %) ont été les grandes gagnantes.

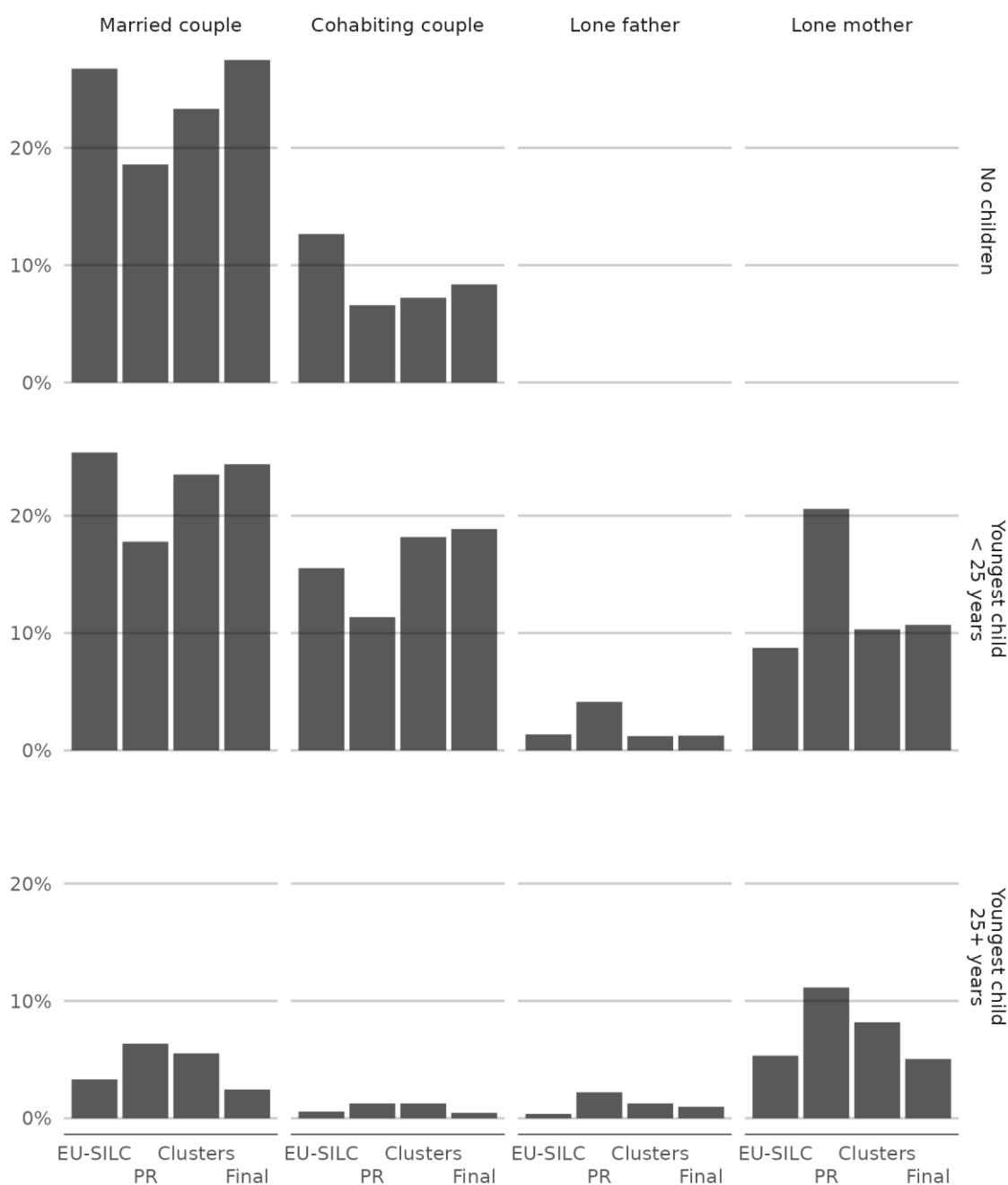
23. L'emploi de la méthode fondée sur les graphes a été motivé par la volonté d'améliorer les statistiques sur les ménages et les familles. Dans la figure 2, nous analysons la distribution du type de noyau familial en fonction de différentes sources. Comme référence, nous utilisons les données des statistiques de l'Union européenne sur le revenu et les conditions de vie de 2022 (EU-SILC 2022), collectées 1 à 5 mois après le moment du dénombrement. Ces données constituent la base des statistiques annuelles sur les ménages et les familles. Les données du registre de la population représentent les familles en fonction du lieu de résidence enregistré. Dans ce registre, comme dans le recensement pilote, la part des familles de partenaires est sous-estimée et celle des familles monoparentales est surestimée. Lorsqu'on procède à la détection des communautés, le nombre de chaque type de couple augmente, bien que l'augmentation soit modeste pour les couples de concubins sans enfants. Enfin, à l'issue du traitement a posteriori, nous obtenons une distribution qui correspond bien aux données EU-SILC. Nous observons encore des disparités parmi les couples en concubinage : la méthode fondée sur les graphes semble sous-estimer la part des couples sans enfants (12,7 % selon les données EU-SILC contre 8,4 % avec les graphes) et amplifier la proportion des familles ayant des enfants en bas âge (15,5 % contre 18,9 %). Néanmoins, nous



considérons que les résultats fondés sur les graphes constituent une amélioration notable par rapport à la représentation initiale des familles par le registre de la population.

24. Depuis la pandémie de COVID-19, le mode opératoire de l'enquête EU-SILC et de l'enquête sur la population active est passé des entretiens en face à face aux entretiens téléphoniques et à l'Internet, ce qui peut affaiblir dans ces enquêtes la fiabilité des données relatives au lieu de résidence. L'utilisation des données d'enquête présente un autre inconvénient : nous observons les ménages en tant qu'unités économiques et non en tant que ménages fondés sur une adresse. Bien que ces unités soient généralement identiques, nous devons reconnaître que les données d'enquête ne rendent pas parfaitement compte des ménages fondés sur une adresse.

Figure 2  
Distribution du type de noyau familial selon différentes sources



Sources : EU-SILC 2022, PR (registre de la population), Clusters (grappes) – méthode fondée sur les graphes, après détection des communautés, Final – méthode fondée sur les graphes, après traitement a posteriori.

## IV. Conclusion

25. L'imprécision du registre estonien de la population en ce qui concerne le lieu de résidence a une influence sur la composition des ménages. Les données du registre ont tendance à privilégier les familles monoparentales.

26. L'exploitation d'autres sources de données administratives peut réduire le biais des statistiques sur les ménages et les familles. Nous avons considéré les personnes et les logements comme les sommets d'un graphe ; les arêtes sont les liens qui apparaissent dans les registres (par exemple, le mariage relie les conjoints, la propriété relie l'appartement à son propriétaire). Un ménage et son logement peuvent alors être considérés comme un graphe secondaire densément connecté ou, en d'autres termes, comme une communauté. Pour trouver ces connexions, nous avons appliqué la méthode de la détection des communautés.

27. Les statistiques obtenues correspondent bien aux estimations de l'UE-SILC concernant les familles et constituent un progrès notable par rapport aux statistiques fondées sur le registre de la population.

## Références

- Blondel, V. D., Guillaume, J.-L., Lambiotte, R. & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10), P10008. <https://doi.org/10.1088/1742-5468/2008/10/P10008>.
- Gale, D. & Shapley, L. S. (1962). College Admissions and the Stability of Marriage. *The American Mathematical Monthly*, 69(1), 9. <https://doi.org/10.2307/2312726>.
- Gortfelder, M. & Puur, A. (2021). Tegelik ja registripõhise elukoha lahknevus ning selle põhjused: 2020. Aasta Eesti tööjõu-uuringu analüüs (lk 31). [https://sisu.ut.ee/sites/default/files/mobiilneelu/files/tp1\\_tlu\\_tegelik\\_ja\\_registripohise\\_elukoha\\_kattuvus\\_ning\\_selle\\_pohjused\\_etu2020\\_analuus\\_gortfelderpuut2021\\_0.pdf](https://sisu.ut.ee/sites/default/files/mobiilneelu/files/tp1_tlu_tegelik_ja_registripohise_elukoha_kattuvus_ning_selle_pohjused_etu2020_analuus_gortfelderpuut2021_0.pdf).
- Porter, M. A., Onnela, J.-P. & Mucha, P. J. (2009). Communities in Networks. *Notices of the American Mathematical Society*, 56(9), 1082–1097.
- Rosvall, M. & Bergstrom, C. T. (2008). Maps of random walks on complex networks reveal community structure. *Proceedings of the National Academy of Sciences*, 105(4), 1118–1123. <https://doi.org/10.1073/pnas.0706851105>.
- Tiit, E.-M., Visk, H., Maasing, E., Levenko, V. & Lehto, K. (2021). Järjekordne rahva ja eluruumide loendus: Milleks ja kuidas? *Akadeemia*, 2021(11), 2009–2064.
- Äär, H. (2017). Coincidence of actual place of residence with Population Register records. *Quarterly Bulletin of Statistics Estonia*, 1, 80–83.
-