



ined
INSTITUT
NATIONAL
D'ÉTUDES
DÉMOGRA
PHIQUES

Les coordonnées de contact dans Fidéli-Nautile : *Quels enseignements en tirer pour la mise en place de protocoles de collecte ?*

Géraldine Charrance, Ined



Travaux menés dans le cadre d'une convention avec l'**Insee**

et présentés lors de l'Assemblée

du 14 octobre 2022, Ined

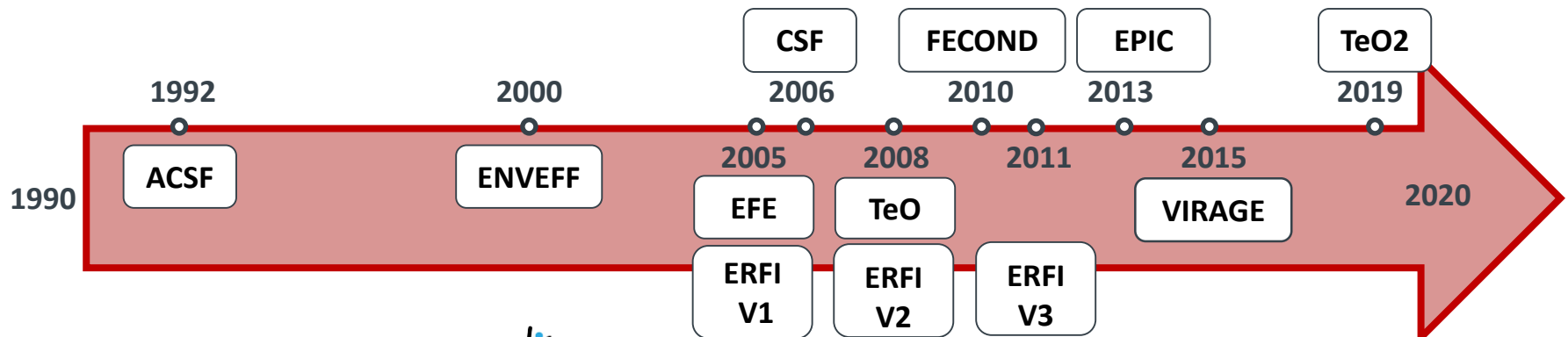
Points abordés

- Contexte des enquêtes à l'Ined
- Analyses sur Nautile 2021 dans le cadre d'une convention avec l'INSEE
 - Coordonnées à l'échelle individuelle
 - Coordonnées mobilisées pour les enquêtes
 - D'autres coordonnées mobilisables ?
- Retours d'expériences sur des enquêtes conduites à partir d'échantillons tirés dans Fidéli
 - EpiCov Vague 1 (Inserm-Drees)
 - Test ERFI 2.0 (Ined)

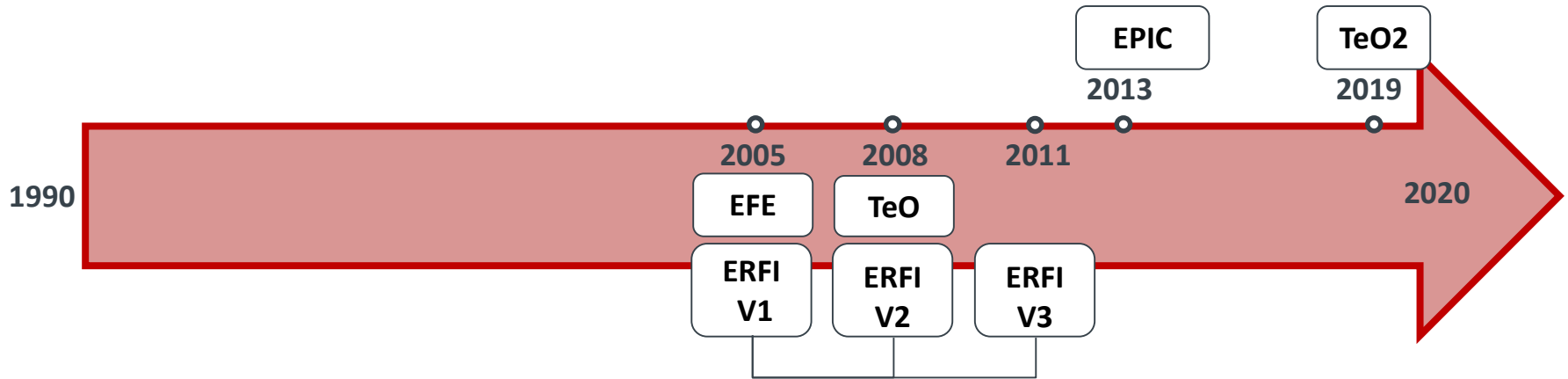
Les enquêtes en population générale

- Population d'intérêt n'a pas de caractéristiques particulières hormis éventuellement :
 - Résidence en France métropolitaine
 - En ménages ordinaires
 - S'inscrivant dans une certaine classe d'âge (champ large)

10 enquêtes Ined répondent à ces critères sur ces trente dernières années (*liste possiblement non exhaustive*)



Les enquêtes en face-à-face



Enquête Familles et Employeurs & Étude des Relations Familiales et Intergénérationnelles:
Echantillon-Maître du recensement de 1999 + Base des logements neufs

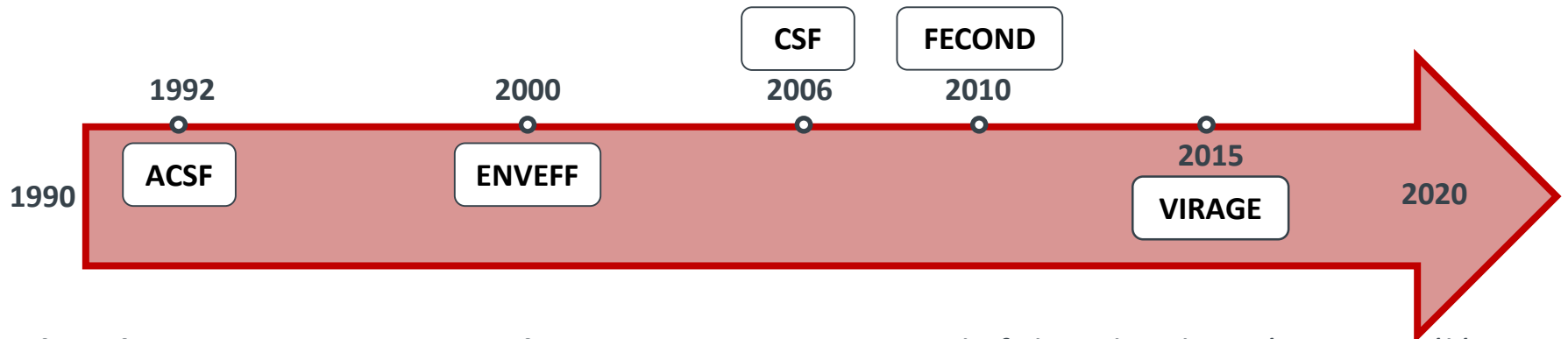
Trajectoires et Origines : EAR 2007 (Enquête Annuelle de Recensement)

Etudes des Parcours Individuels et Conjugaux : EAR 2012

Trajectoires et Origines 2 : EAR 2018

Enquêtes menées avec le concours de l'INSEE, en mobilisant son réseau d'enquêteurs

Les enquêtes par téléphone



Analyse des Comportements Sexuels en France : Tirage au sein du fichier des abonnés France Télécom

Enquête Nationale sur les Violences Envers les Femmes en France : Constitution d'une base de sondage à partir du fichier des abonnés FT + incrémentation d'une unité (pour prendre en compte les numéros sur liste rouge)

Contexte de la Sexualité en France : Un échantillon « Fixes » : Constitution d'une base de sondage à partir du fichier des abonnés FT + incrémentation d'une unité (pour numéros liste rouge) + un échantillon « Mobiles », obtenu par génération aléatoire

Fécondité – Contraception – Dysfonctions sexuelles : Génération aléatoire de numéros à préfixe géographique (01 à 05) et de numéros de mobiles (06). Critère de mobile exclusif

Violences et Rapports de Genre : Génération aléatoire de numéros à préfixe géographique (01 à 05), de numéros de box (09) et de numéros de mobiles (06). Critère de mobile exclusif assoupli

Problématiques actuelles

- Enquêtes en face-à-face :
 - Extrêmement coûteuses
 - Envisageables uniquement pour des collectes sur de petites zones géographiques
 - Non recommandées pour aborder des thématiques sensibles
- Enquêtes par téléphone :
 - Nombreuses sollicitations des usagers (démarchage téléphonique)
 - Stratégies de filtrage des numéros inconnus
 - Nécessité de pouvoir informer les personnes de notre démarche en amont des tentatives d'appel pour leur permettre de nous identifier et de s'informer sur l'enquête.

Par génération aléatoire de numéros : possible uniquement pour les personnes retrouvées avec l'annuaire inversé (via un courrier postal) et sur les lignes mobiles (via l'envoi d'un SMS)

Une opportunité en or : Fidéli/Nautile

Fidéli (*Fichier Démographique sur les Logements et les Individus*)

Assemblage raisonné de données administratives pour produire des statistiques démographiques

- ✓ Données d'origine fiscale (taxe d'habitation, propriétés bâties, fiches d'imposition des personnes, fichier des déclarations de revenus)
- ✓ Données complémentaires contextuelles
- ✓ Informations sur les grandes masses de revenus déclarées

Nautile (*Nouvelle Application Utilisée pour le Tirage des Individus et des Logements des Enquêtes*)

Base de sondage constituée à partir de Fidéli

Une opportunité en or : Fidéli/Nautile

Nautile présente de nombreux avantages en théorie :

- Exhaustive
- Sans double compte
- Avec de l'information auxiliaire pouvant être mobilisée :
 - pour échantillonner (principales caractéristiques sociodémographiques, données fiscales)
 - pour réaliser l'enquête (coordonnées de contact : adresse postale, numéro de téléphone, adresse mail)

Cependant, nécessité d'en savoir davantage sur cette base de sondage, notamment sur la présence des coordonnées de contact, l'objectif étant de savoir si elle peut être mobilisée pour la conduite d'enquêtes multimode CATI/CAWI.

Ne faisant pas partie du Service Statistique publique, plusieurs étapes sont requises pour disposer d'un échantillon tiré dans Nautile :

- Obtenir l'avis d'opportunité du CNIS
- Obtenir le label d'intérêt général et de qualité statistique (Comité du Label)
- Obtenir l'accord du Comité de Direction
- Passer au Comité du Secret
- Mettre en place une convention bipartite avec la Direction de la Méthodologie et de la Coordination Statistique et Internationale de l'INSEE

Principaux résultats des analyses conduites sur Nautile 2021

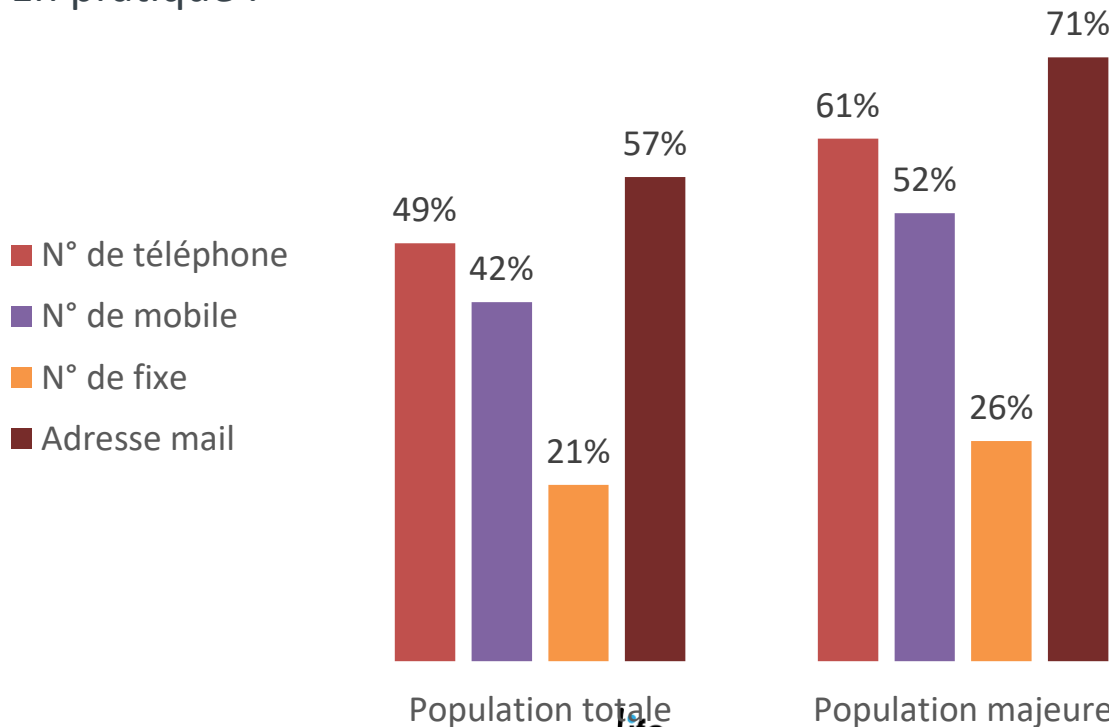
Analyses menées en se restreignant aux individus s'inscrivant dans le champ des enquêtes en France métropolitaine.

Coordonnées de contact à l'échelle individuelle

En théorie :

- Deux champs « Téléphone » (tel_port et tel2) et un champ « Adresse mail » par individu pour les référents principaux & les conjoints.
- Pas de coordonnées de contact pour les personnes à charge.

En pratique :

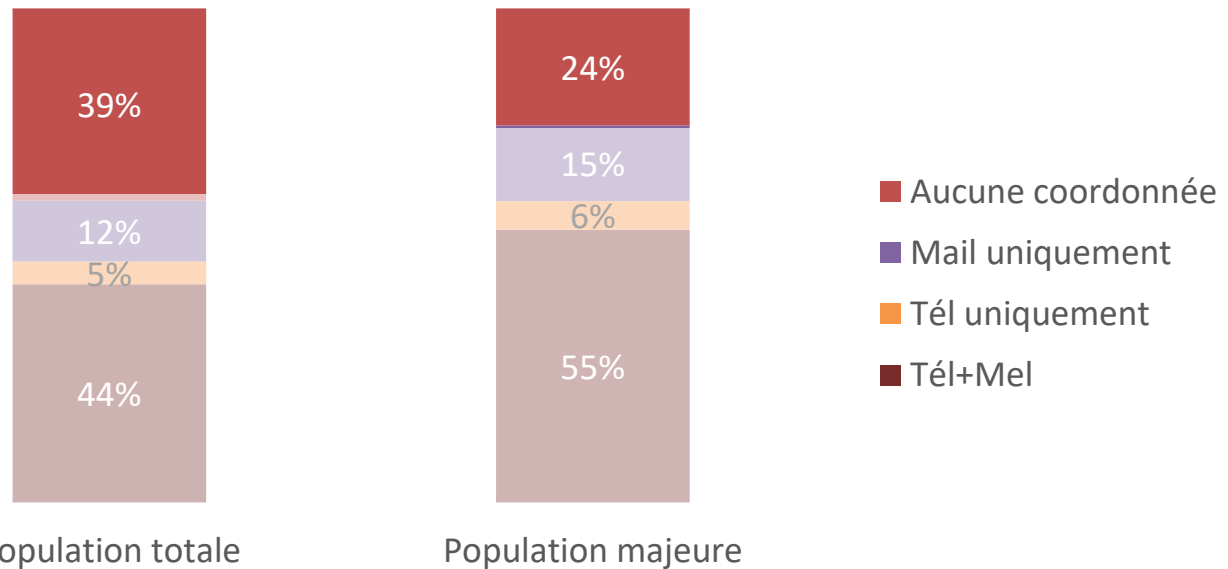


Parmi la population joignable par téléphone :

- ✓ 72% avec un seul numéro (80% de mobile)
- ✓ 28% avec deux numéros (99% fixe + mobile)

Coordonnées de contact à l'échelle individuelle

Répartition de la population selon les coordonnées disponibles



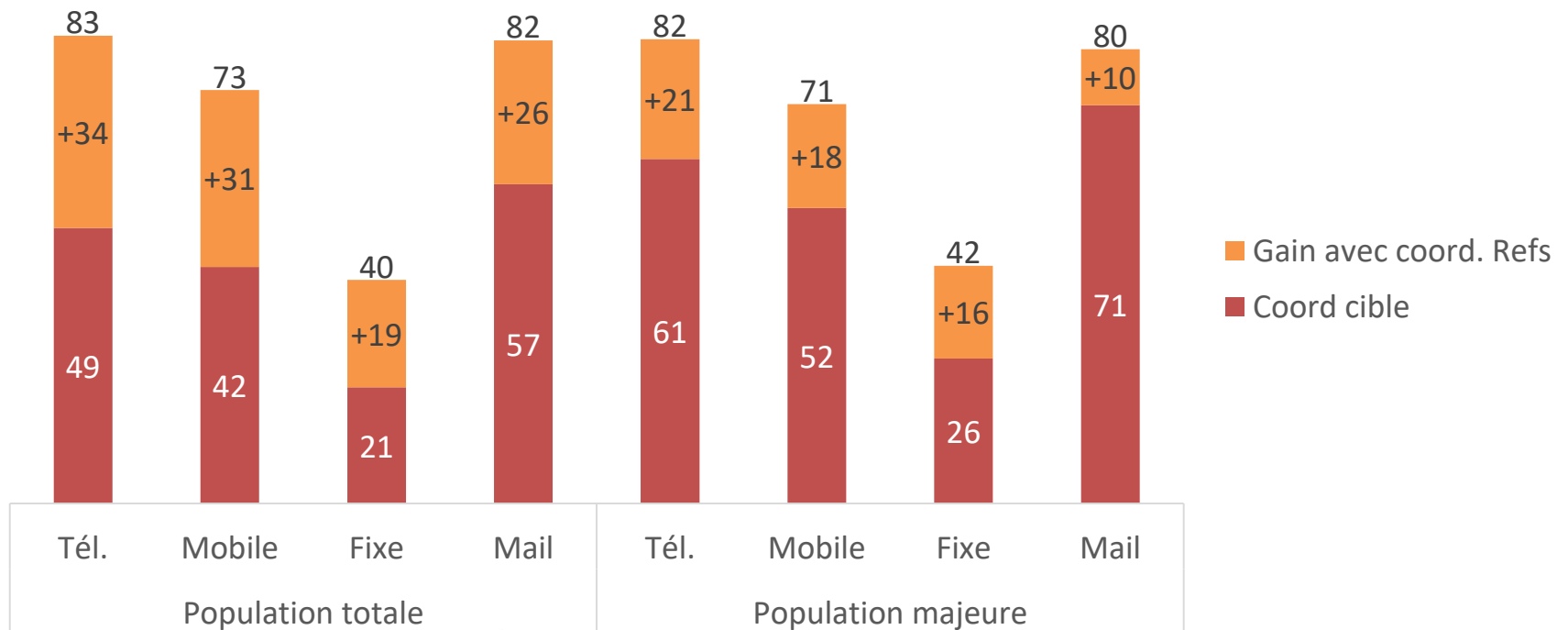
Populations les plus touchées par l'absence de coordonnées de contact :

- Femmes
- Moins de 25 ans
- 75 ans ou plus
- Personnes à charge, enfants (ni référent, ni conjoint)
- Familles monoparentales ou ménage « complexe »
- Personnes vivant en dessous du seuil de pauvreté

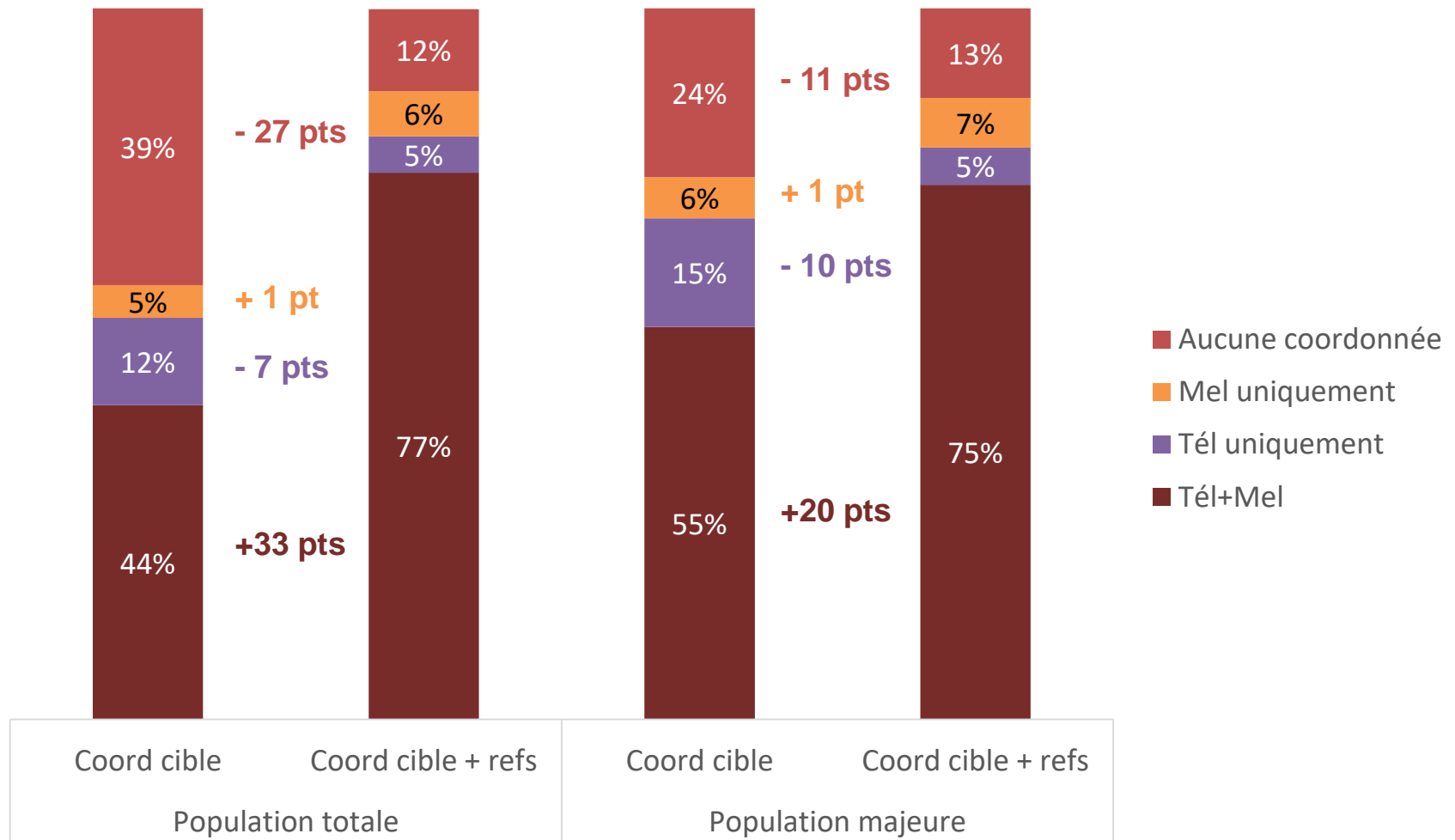
Coordonnées mobilisées pour les enquêtes

Pour la conduite d'enquêtes :

- ✓ Coordonnées de la cible
- ✓ Coordonnées du référent du logement dans lequel la cible réside (et celles de son/sa conjoint·e éventuel·le)



Coordonnées mobilisées pour les enquêtes



Evolution en fonction des caractéristiques sociodémo.

Populations pour qui la prise en compte des coord. Ref engendre à un gain important en termes de coordonnées téléphoniques (population majeure) :

- Femmes
- Moins de 25 ans
- Autre statut que référent fiscal
- Couples avec ou sans enfant

Populations pour qui la prise en compte des coord. Ref engendre à un gain important en termes d'adresses mail (population majeure) :

- Moins de 25 ans
- Personnes à charge (enfants)
- Couples avec enfants (probablement les enfants), familles monoparentales

Populations surreprésentées dans la catégorie « Ni tél. ni mail » en tenant compte des coord ego et des référents logements:

- 20-24 ans et les 75 ans et plus
- Les personnes vivant seules
- Les personnes en dessous du seuil de pauvreté

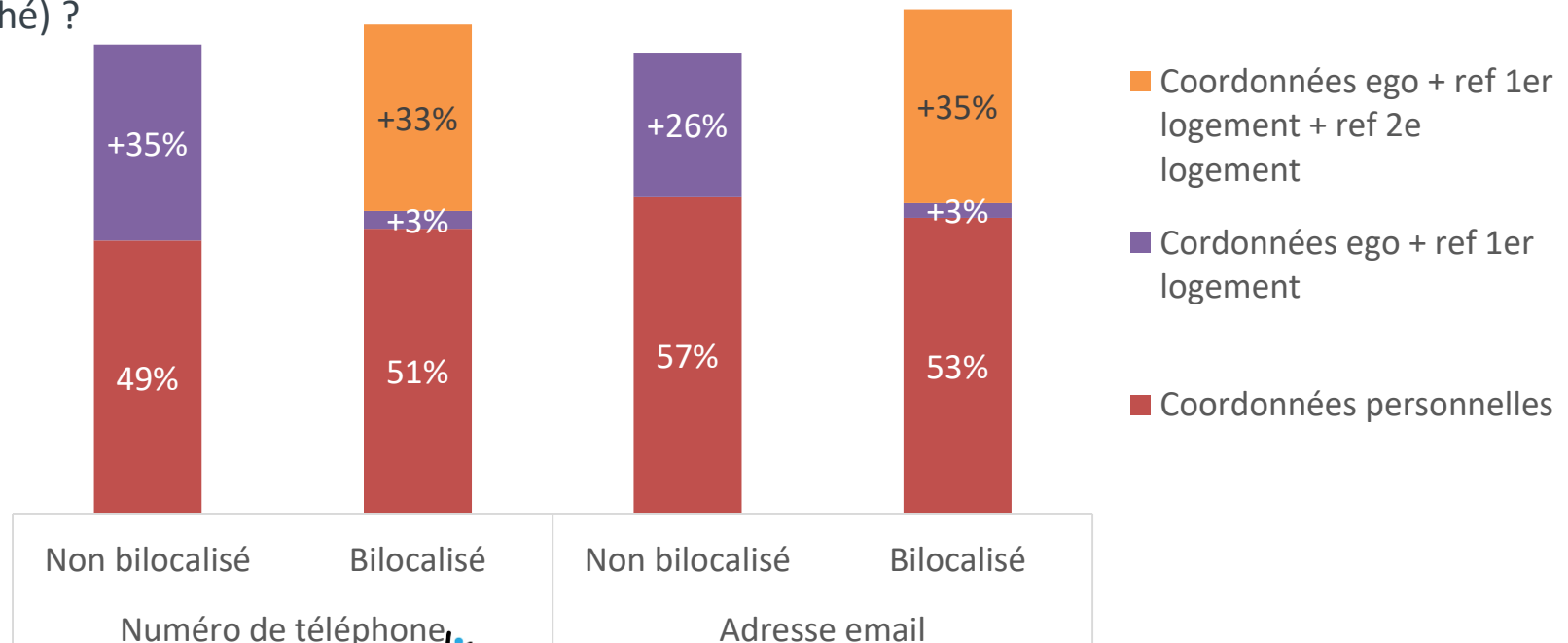
D'autres coordonnées mobilisables ?

Cas des individus bilocalisés

Combien sont-ils ? 1,6% de la population totale et 2,0% de la population majeure.

Qui sont-ils ? Principalement des jeunes vivant seuls mais encore rattachés au foyer fiscal de leur(s) parent(s) (ne résident donc pas avec les référents de leur foyer fiscal)

Peut-on envisager de mobiliser les coordonnées de contacts du/des référents du second logement (logement dans lequel résident les référents du foyer fiscal auquel le jeune est rattaché) ?



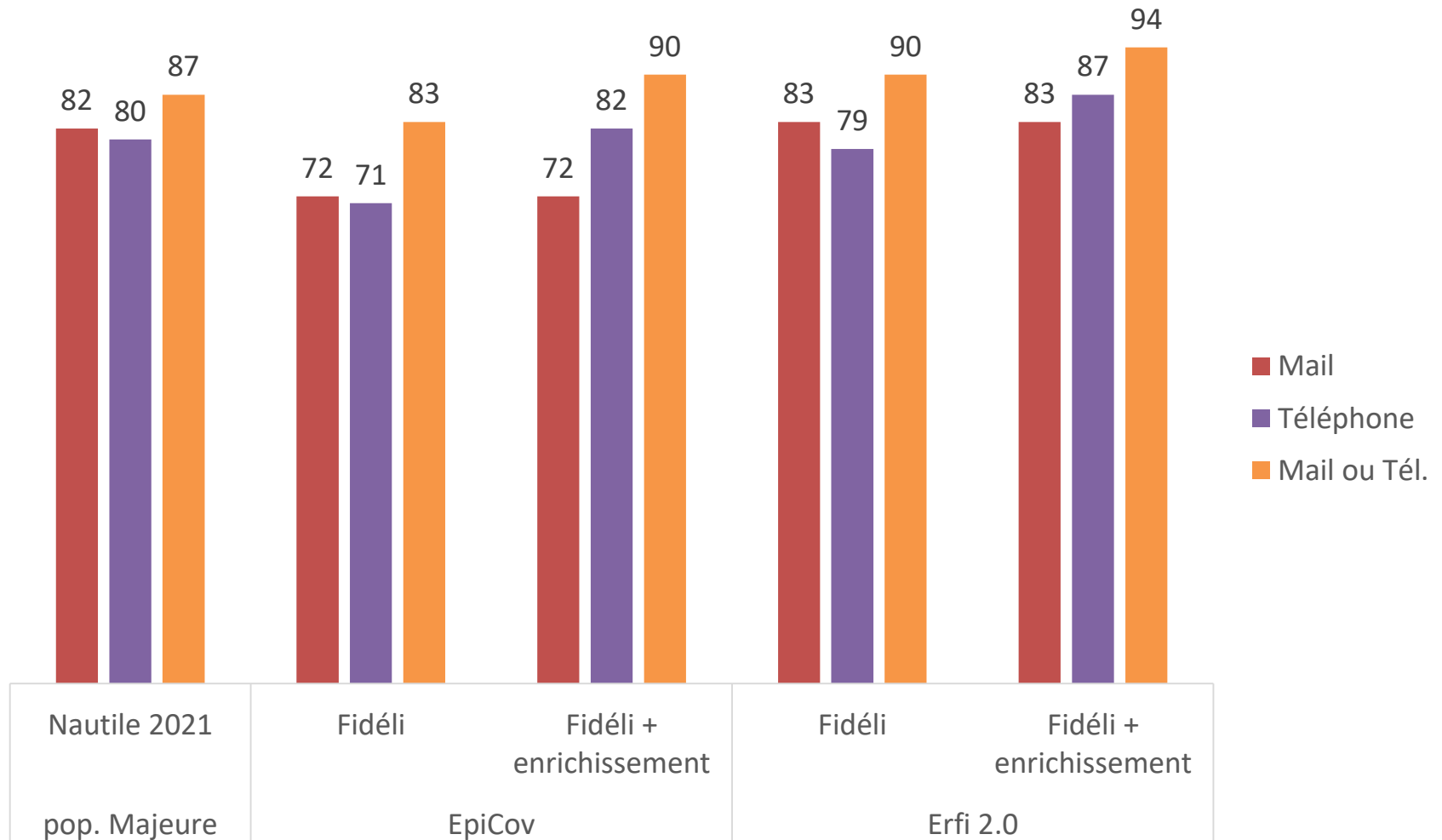
Retours d'expériences sur des enquêtes menées sur la base d'échantillons tirés dans Fidéli

Retours d'expériences – EpiCov (Vague 1) et test Erfi 2.0

	EpiCov (V1)	Test Erfi 2.0
Collecte	Mai 2020	Nov 2021-Février 2022
Plan de sondage	Surreprésentation des ménages « pauvres »	Stratification proportionnelle sur la typologie de coordonnées de contact
Taille de l'échantillon	371 000 Fr Entière, dont 350 000 en métropole	3199
Champ	15 ans ou plus	18-79 ans
Modes de contact	Protocole d'annonce et de relance particulièrement fourni, mobilisant toutes les coordonnées à disposition (adresse postale, mails, numéros de téléphone)	
Modes de collecte	<p>→ 4 lots soumis à des protocoles multimodes (internet/téléphone) :</p> <ul style="list-style-type: none"> • 3 lots selon un multimodes concurrentiel (deux modes ouverts simultanément) • 1 lot selon un multimode concurrentiel différé (l'exploitation téléphonique n'a débuté que 2 semaines après l'ouverture du lot) <p>→ 16 lots soumis à un protocole CAWI exclusif</p>	<p>→ 4 lots soumis à un protocole multimodes séquentiel CATI-CAWI (avec une déclinaison d'incitations financières)</p> <p>→ 4 lots soumis à un protocole séquentiel CAWI-CATI (avec une déclinaison d'incitations financières)</p> <p>→ 1 lot soumis à un protocole CAWI exclusif</p>

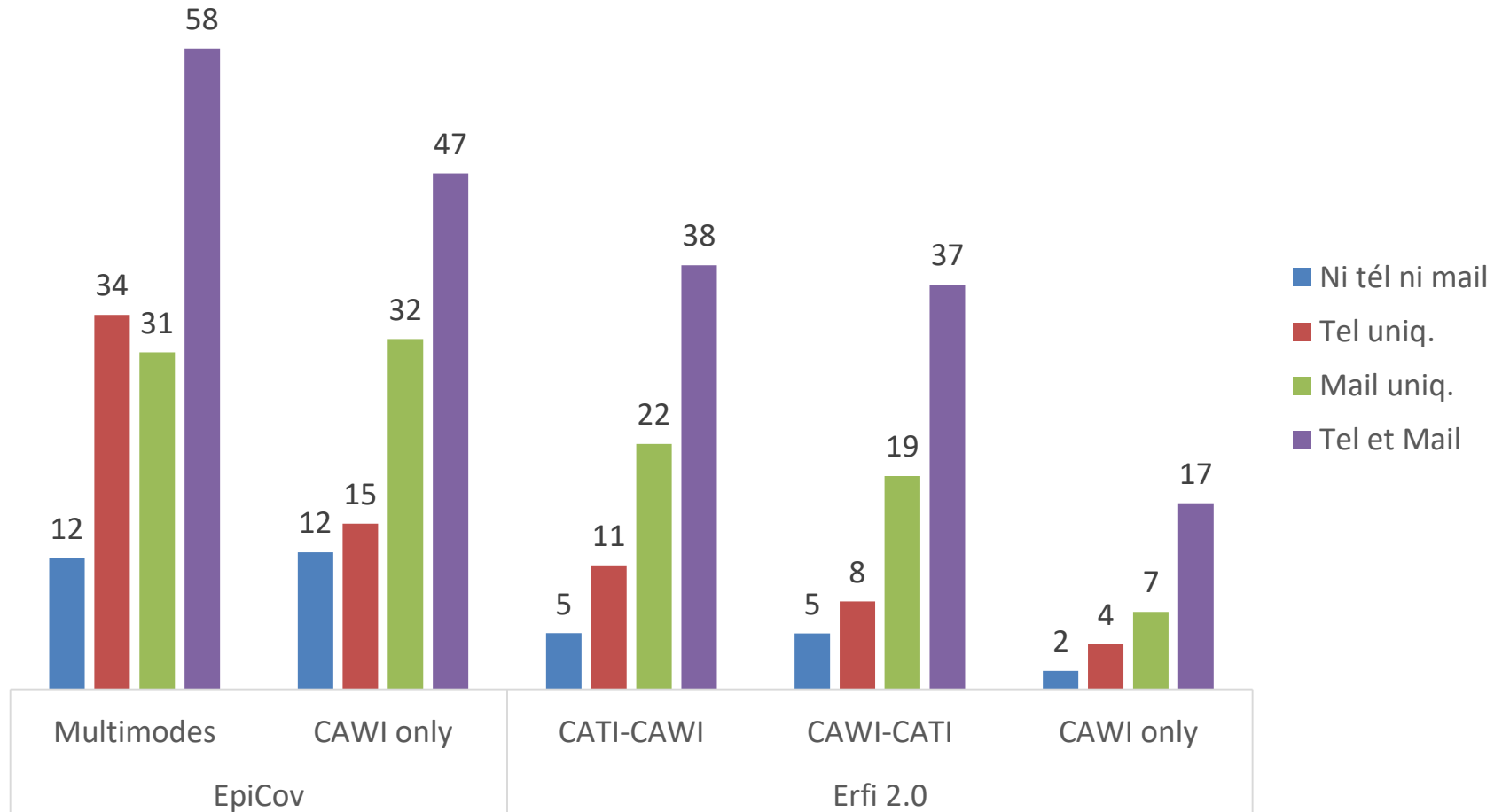
Coordonnées disponibles dans Fidéli et enrichissement

EpiCov et Erfi 2.0



Taux de réponse selon protocole et coord disponibles

Epicov et Erfi 2.0



Courrier Postal

- Ne pas recevoir de courrier postal semble très pénalisant (-20 pts sur la proba de rép.)
- L'absence de courrier postal ne semble pas pouvoir être facilement compensée par un mail ou un SMS

Type et nombre de modes de contact

- Relances par mail plus efficaces que par SMS (et moins coûteuses)
- La pluralité des modes de contact est associé à une probabilité de réponse supérieure (surtout vrai pour les coord cible mais vérifiée également pour les coord de tiers).

Coord cible vs coord tiers

Effet coord cible vs tiers plus marqué sur les relances mail que par les relances SMS. Pas de diff. significative sur la proba de réponse selon si le SMS est envoyé à la cible ou à un tiers.

Collecte par téléphone : Mieux cibler pour limiter l'effort de collecte

Epicov

		Prob qu'un appel aboutisse à un décroché	Prob qu'un appel décroché aboutisse à une interview
Proba situation de ref		16,9	25,7
Fidéli	Fixe géo	<i>Ref.</i>	<i>Ref.</i>
	Mobile cible	+4,4 pts	+9,5 pts
	Mobile Autre	+1,8 pts	-3,6 pts
	Box	-5,8 pts	ns
Annuaire	Fixe géo	-1,3 pts	-5,5 pts
	Mobile	-3,8 pts	ns
	Box	-7,8 pts	ns
Numéro terrain		+12 pts	+21,5 pts

Numéros annuaire : répondent plutôt par téléphone -> Plus-value de l'enrichissement
 Effet négatif du PND présent également sur la collecte par téléphone, mais moins fort que sur la collecte Internet

Conclusion

- Fidéli/Nautile est une source pertinente au vu des analyses menées sur Nautile 2021 et les premiers retours d'expériences
- Il est essentiel et indispensable de mobiliser les coordonnées d'autres personnes
- Piste déjà évoquée à creuser : Comment enquêter les « sans téléphone ni mail »
 - Récupérer d'autres coordonnées via Fidéli : Piste du deuxième logement (individus bilocalisés) surtout sur des enquêtes visant des jeunes
 - Inciter les personnes à se manifester pour fournir des coordonnées (dans test Erfi 2.0, 13% des personnes sans ni tel ni mail se sont manifestées pour communiquer des coordonnées téléphoniques)
- Avec Erfi 2.1 et FamEmp, de nouvelles analyses sur de larges échantillons issus de Fidéli pourront être menées.

Merci pour votre attention

Contact : geraldine.charrance@ined.fr

Pour en savoir plus sur les résultats de la 1^{ère} vague d'EpiCov :

http://www.jms-insee.fr/2022/S20_3_ACTE_CHARRANCE_JMS2022.pdf