

# ENQUÊTES TELEPHONIQUES AUPRES DES MÉNAGES : PEUT-ON UTILISER L'ANNUAIRE ?

*Jérôme ACCARDO, Emmanuelle BASCHERI, Laurence JALUZOT*

*Insee, Département des Prix, des Ressources  
et des Conditions de vie des Ménages*

## Introduction

Si la plupart des enquêtes de l'Insee auprès des ménages sont réalisées en face à face, certaines sont faites par téléphone. C'est notamment le cas de l'enquête mensuelle de conjoncture CAMME et de l'enquête annuelle sur les Technologies de l'Information et de la Communication (TIC). D'une part leur questionnaire est léger et se prête à une passation par téléphone. D'autre part, en raison de leur fréquence, il serait trop coûteux de les réaliser en face à face.

On est alors confronté à l'inadéquation de la base de sondage habituelle de l'Insee, l'échantillon-maître issu du recensement, puisque cette base ne comporte pas de numéros de téléphone. La seule solution praticable est de recourir à l'annuaire<sup>1</sup>, autrement dit à la liste des numéros de téléphone que l'opérateur dit "historique" (France Telecom) est requis de maintenir et de tenir à la disposition d'annuairistes. Ceux-ci réalisent à la demande des extractions pour le compte de leurs clients, en général pour des opérations de prospection commerciale.

En tant que base de sondage, l'annuaire souffre d'un défaut de couverture de la population. Il y a quatre motifs d'absence de la base :

- le ménage n'a pas de téléphone : ce cas est très rare.
- le ménage n'a pas de ligne fixe mais dispose d'au moins un téléphone portable. On le dira monoéquipé de mobile (MEM). On les estime actuellement à environ 13%.
- le ménage est abonné à une ligne fixe France télécom, mais n'est pas joignable (liste rouge ou orange<sup>2</sup>). On compte environ 12% de tels ménages<sup>3</sup>.
- le ménage a une ligne fixe, mais chez un autre opérateur (les "dégrouvés"). La proportion de ces ménages a rapidement augmenté ces dernières années : en 2005, 34% des individus de 18 ans et plus disposant d'une ligne fixe étaient clients d'un opérateur autre que France Télécom, mais seule une faible minorité avait complètement abandonné FT (les "dégrouvés totaux"). En 2008, on compte 43% de dégroupés, dont les 2/3 sont des dégroupés totaux (Bigot, Croutte, 2008). Le dégroupage, même total, n'entraîne pas nécessairement la disparition de l'annuaire<sup>4</sup>. Dans cette contribution, on conviendra néanmoins d'appeler "dégrouvé" les ménages ou individus ne figurant pas dans l'annuaire parce qu'ils ne sont plus chez France Télécom.

Au total, désormais près de 45% des ménages ne figurent pas dans l'annuaire. Ce chiffre, tiré de l'enquête TIC 2008, est confirmé par les expériences de recherches de numéros téléphoniques à partir d'adresse conduites dans le cadre de l'enquête "Changements Organisationnels et Informatisation (COI) 2006" ou pour l'enquête "Vie Quotidienne et Santé (VQS) 2007".

---

<sup>1</sup> C'est un sous-ensemble de l'annuaire papier ou en ligne (voir note suivante).

<sup>2</sup> Les numéros en liste orange peuvent apparaître dans l'annuaire papier ou en ligne, mais ne peuvent être sélectionnés par l'annuairiste pour des opérations de sondage, démarchage, etc.

<sup>3</sup> Ces estimations sont issues des enquêtes CVS 2007 et 2008 et TIC 2005.

<sup>4</sup> Cependant, les deux enquêtes analysées ici excluent, pour des raisons liées à la chaîne d'échantillonnage, les numéros dégroupés (i.e. ceux dont le préfixe est 07, 08 ou 09) qui se trouvent dans l'annuaire.

Tableau 1 : Equipement téléphonique des ménages et présence dans l'annuaire (%)

	2008		2005 (observé)	2005 (ventilation estimée)	2008
Avec une ligne fixe	87	Annuaire	70	63	56
		Dégroupé-HA hors liste		7	14
		France Télécom (*) sur liste	16	14	12
		Dégroupé-HA sur liste		2	5
Sans ligne fixe	13	Mono-équipés de mobile	13	13	12
		Sans téléphone	1	1	1

Source : Insee, enquêtes TIC 2008 et CVS 2005.

Champ : ensemble des ménages ordinaires, France métropolitaine.

Lecture : 87% des ménages disposent d'une ligne fixe. 56% figurent dans l'annuaire.

Note : Dégroupé-HA = dégroupé hors annuaire

(\*) éventuellement avec d'autres opérateurs pour un autre numéro.

La comparaison avec l'enquête de victimation "Cadre de vie et sécurité (CVS) 2005" suggère une dégradation de la couverture de la base. Plus précisément : l'enquête CVS 2005 ne permettant pas d'identifier les dégroupés, on fait l'hypothèse que le poids des ménages en liste rouge est resté constant. Dans ces conditions, le poids des ménages dans l'annuaire et des dégroupés hors liste était de 70% environ (tableau 1). En admettant qu'en juin 2005, les dégroupés représentaient le dixième de cet ensemble (voir Bigot, Croutte, 2008, p.24), on peut estimer que la base a perdu de l'ordre de 7 % des ménages en trois ans.

Le défaut de couverture de l'annuaire n'est pas en soi rédhibitoire (l'échantillon maître lui aussi ne comprend qu'une part des ménages en France). En revanche il soulève deux questions :

- l'annuaire est-il représentatif de la population ?
- Conduit-il à des estimations non biaisées ?

## 1. Une base non représentative

La réponse à la première question est clairement non :

- plus un ménage est jeune, moins il est joignable. 57% des plus de 60 ans vivent dans un ménage inscrit dans l'annuaire<sup>5</sup>. Mais dans les tranches plus jeunes, ils n'y figurent pas en majorité (le taux s'abaissant jusqu'à 32 % chez les 20-29 ans). Chez les plus jeunes, l'absence de ligne fixe est relativement répandue (17% des individus contre 9% chez les 40-59 ans).
- Les agriculteurs sont les plus joignables : 70% d'entre eux vivent dans un ménage qui figurent dans l'annuaire. Les moins joignables sont les cadres (32% dans l'annuaire) bien qu'ils disposent à peu près tous d'une ligne fixe (96%). Le taux de monoéquipés de mobile est particulièrement élevé chez les ouvriers (16%, contre 10% pour l'ensemble). Et il atteint 20% chez les chômeurs.

L'examen du niveau de revenu ou de diplôme confirme le diagnostic. Etre hors annuaire renvoie à deux types de population :

- Les possesseurs d'une ligne fixe hors annuaire (les "dégroupés") sont des individus relativement plus jeunes, plus aisés et plus diplômés que la moyenne. Ils sont plus souvent actifs occupés.
- Les mono-équipés de mobile sont plus jeunes, plus modestes et moins diplômés que les autres. Les chômeurs y sont sur-représentés.

Les personnes sur liste d'exclusion<sup>6</sup> sont un peu plus âgés que ceux dans l'annuaire et sensiblement plus aisés. Ils en sont en revanche assez proche du point de vue de la situation vis à vis de l'activité et du point de vue des diplômes.

<sup>5</sup> Source : Insee, enquête TIC 2008

<sup>6</sup> ou plus exactement dans des ménages sur liste (mais non dégroupés).

Tableau 2 : Comparaison des individus selon leur équipement téléphonique<sup>7</sup>.

	Ensemble	Annuaire	Sur liste	Dégroupé	MEM
Part dans la population (%)	99	56	11	22	10
Age	48,8	51,7	54,4	41,8	42
Revenu mensuel du ménage auquel l'individu appartient (en euros)	2 311	2 254	2719	2 481	1 775
Sans diplôme ou niveau primaire (%)	62	64	62	49	73
Diplôme de niveau secondaire (%)	16	16	15	19	8
Diplôme du supérieur (%)	23	19	23	32	19
En emploi (%)	50	45	44	67	50
Au chômage (%)	5	4	5	6	9

Source : Insee, enquête TIC 2008

Champ : Individus de 18 ans et plus, France métropolitaine.

Lecture : 56% des individus vivent dans un ménage figurant dans l'annuaire. Leur âge est en moyenne de 51,7 ans. 64% possèdent au plus un diplôme de niveau primaire.

Si la population de l'annuaire n'est pas représentative, faut-il en conclure que les résultats des enquêtes téléphoniques sont biaisés ? Ou, plus exactement, le redressement de la structure socio-démographique des échantillons issus de l'annuaire élimine-t-il la source de biais sur les résultats ? La réponse est ici plus délicate. D'autant qu'elle dépend, évidemment, de la nature des variables étudiées.

Pour les enquêtes de conjoncture, le défaut de couverture, parfaitement reconnu dès les débuts de l'enquête mensuelle de conjoncture a toujours été considéré comme négligeable. Jusqu'au début des années 2000, cette position pouvait être admise en première approximation. Jusqu'à récemment, en effet, le défaut de couverture se limitait soit aux ménages sans téléphone (très peu nombreux), soit à ceux en liste d'exclusion dont on ne voyait pas pourquoi ils auraient eu, de ce seul fait, des opinions très différentes de ménages de même caractéristiques mais figurant dans l'annuaire.

Au cours des années récentes, néanmoins, l'érosion de la base s'est accentuée avec l'apparition des monoéquipés mobiles et des dégroupés. Avec un ménage sur deux joignables, on peut légitimement craindre que le défaut de couverture soit devenu réellement gênant. Et ce qui est vrai pour l'enquête de conjoncture l'est a fortiori pour l'enquête TIC : dans cette enquête consacrée à l'équipement et aux pratiques des individus en informatique et téléphonie, il faut considérer *a priori* qu'être hors de l'annuaire a une incidence sur les réponses.

Malheureusement, pour l'instant il n'existe pas d'alternative. D'une part l'annuaire universel souhaité par le régulateur n'est toujours pas en vue. D'autre part, les enrichissements de la base proposés par les annuairistes, notamment à partir de listes d'abonnés à d'autres opérateurs, ou à partir de fichiers de numéros de portables sont peu contrôlés (du point de vue statistique) et peuvent tout aussi bien aboutir à une dégradation de la qualité de la base plutôt qu'à une amélioration (par exemple les fichiers de numéros de portables que proposent les marchands de listes proviennent souvent de grands magasins qui récupèrent les numéros indiqués par les clients lorsqu'ils s'inscrivent à une carte de fidélité ; le caractère d'emblée biaisé de ces échantillons est évident. En outre leur utilisation avec d'autres bases suppose d'avoir résolu la question des doublons éventuels).

Enfin il faut exclure tout recours à une procédure de génération aléatoire de numéros de téléphone (Roy, Vanheuverzwyn, 2003), entre autres raisons parce qu'une enquête de l'Insee prévoit toujours l'envoi préalable à l'enquêté d'une "lettre-avis" officielle l'informant de sa sélection dans l'échantillon.

Dans ces conditions, plusieurs études ont été réalisées depuis 2006 dans la division Conditions de vie des ménages pour évaluer l'impact de ce défaut de couverture sur les résultats. Elles s'appuient sur deux opérations méthodologiques réalisées sur les enquêtes TIC et CAMME en octobre 2005 et en

<sup>7</sup> Les chiffres relatifs au très petit nombre des individus dans des ménages sans téléphone ne sont pas significatifs. Ils ne sont pas rapportés ici.

avril 2008. Dans les deux cas, le principe a été le même : on a comparé les résultats d'enquêtes téléphoniques, donc sur des échantillons issus de l'annuaire, avec ceux d'enquêtes conduites auprès d'un échantillon issu de l'échantillon-maître.

Les études conduites concernent aussi bien l'impact sur les résultats des enquêtes TIC et CAMME. Dans le cadre limité de cette contribution, on ne présentera ici que ce qui concerne les TIC. L'impact sur les résultats collectés dans l'enquête de conjoncture est renvoyé en annexe.

## **2. La mesure du biais sur les résultats induit par le défaut de couverture**

### **2.1 Les opérations méthodologiques d'octobre 2005 et d'avril 2008.**

En octobre 2005 l'Insee a conduit une enquête PCV auprès d'environ 5600 ménages répondants sur les TIC. Il s'agissait donc, comme toutes les enquêtes de ce dispositif, d'une enquête en face à face auprès d'un échantillon de ménages issu de l'échantillon-maître. Du questionnaire, on a tiré dix questions qu'on a introduites dans l'enquête de conjoncture du même mois, réalisée par téléphone auprès de 2 000 ménages répondants. On a donc pu comparer les réponses obtenues pour ces dix questions sur les deux types d'échantillon.

Pour l'enquête d'avril-mai 2008, une double procédure d'échantillonnage a été mise en place. L'enquête TIC d'avril 2008 a été conduite par téléphone auprès d'un échantillon de 7 000 ménages<sup>8</sup> tiré dans l'annuaire téléphonique et d'un échantillon de 4 000 ménages tiré dans l'enquête annuelle du recensement (EAR) 2007<sup>9</sup>. De ce second échantillon on a écarté, après recherche manuelle, les ménages figurant dans l'annuaire. Il est resté 2 000 fiches-adresse qui ont été enquêtées en face à face<sup>10</sup>. Pour des raisons d'organisation et d'optimisation des coûts, l'enquête complémentaire n'a été réalisée que dans 4 directions régionales de l'INSEE : Provence-Alpes-Côte d'Azur, Ile-de-France, Lorraine, Limousin.

Le questionnaire passé aux deux échantillons était, sur la partie TIC, le même. On compte, pour l'échantillon téléphonique (c'est à dire les ménages figurant dans l'annuaire) 4 791 répondants et pour l'échantillon issu de l'EAR, 1 287 répondants hors annuaire.

Il faut souligner une difficulté importante. L'enquête TIC 2005 ne distinguait que deux catégories de ménages : ceux disposant d'une ligne fixe et les monoéquipés en mobile. On ne peut rien en tirer sur les écarts existants, parmi les possesseurs d'une ligne fixe, entre ceux dans l'annuaire et ceux qui n'y figurent pas (en liste rouge ou dégroupés). Ce point a deux conséquences :

- l'analyse de l'évolution, entre 2005 et 2008, du biais observé sur la mesure des équipements et des comportements TIC doit se limiter à l'écart entre les possesseurs de fixe (dans l'annuaire ou non) et les monoéquipés en mobile.
- le redressement de l'enquête téléphonique TIC de 2007, ne pouvant se fonder que sur cette information limitée, n'a que partiellement corrigé les biais induits par le défaut de couverture de l'annuaire (voir section 3 ci-après).

### **2.2 L'effet de l'équipement téléphonique sur l'équipement et l'usage des TIC**

L'analyse des évolutions entre TIC 2005 et TIC 2008 doit donc se restreindre à la comparaison entre les abonnés au téléphone fixe et les sans téléphone fixe (essentiellement les monoéquipés de mobile, MEM). En outre, elle doit se limiter à l'équipement (en ordinateur et en accès Internet), les questions sur les comportements ayant trop changé entre les deux enquêtes<sup>11</sup> pour être valablement rapprochées dans le contexte de cette étude méthodologique.

<sup>8</sup> Plus précisément, des numéros de téléphones correspondant à une adresse unique. Les abonnés possédant plusieurs lignes ne figurent donc qu'une seule fois.

<sup>9</sup> L'échantillonnage dans le dernier recensement disponible était motivé par le besoin de disposer d'adresses suffisamment récentes pour qu'une recherche dans l'annuaire soit pertinente.

<sup>10</sup> Dans une cinquantaine de cas, les enquêteurs ont, après une première visite, récupéré un numéro de téléphone et conduit l'enquête par téléphone.

<sup>11</sup> Le questionnaire de l'enquête TIC 2005 avait été élaboré avant l'adoption du questionnaire retenu depuis par Eurostat. En particulier il utilisait des périodes de référence de durées différentes.

Malgré tout, il apparaît clairement que l'écart entre les deux populations s'est accentué. Plus précisément, chez les possesseurs d'une ligne fixe, l'équipement en ordinateur ou en accès à Internet, déjà supérieur en 2005, a fortement augmenté. Il a en revanche stagné chez les MEM (tableau 3). On pourrait alors penser qu'une enquête dans l'annuaire seulement donnera en 2008 des résultats plus biaisés qu'en 2005.

Tableau 3 : Evolution de trois indicateurs TIC selon l'équipement du ménage

	2005		2008	
	Téléphone fixe	MEM et sans tél	Téléphone fixe	MEM et sans tél
Ménages équipés d'un ordinateur	60	39	79	38
Ménages ayant accès à Internet	47	10	75	15
Ménages ayant accès au haut débit	35	8	65	11

Source : Insee, enquêtes TIC 2005 et TIC 2008

Champ : ménages comprenant un individu âgé de 16 à 74 ans, France métropolitaine.

Lecture : en 2005, 60% des ménages équipés d'un téléphone fixe ont un ordinateur chez eux.

La situation est en réalité plus compliquée ; en effet, la population des possesseurs de ligne fixe qui, du point de vue de l'équipement et des comportements en TIC, se différencie fortement des sans ligne fixe, est elle-même hétérogène. Le fait d'être ou non en liste rouge a peu d'importance, pour les individus (ou ménages) chez France Télécom. En revanche, être absent de l'annuaire parce que dégroupé est associé à un investissement beaucoup plus fort que les autres dans les TIC (tableau 4).

Une partie de l'écart est certes due aux différences de structure socio-démographiques des diverses populations. Une stratification par âge ou par diplôme peut atténuer fortement les différences à l'intérieur des possesseurs de ligne fixe, entre les dégroupés et les abonnés de France Télécom. Ainsi, l'estimation de modèle de régression logistique (non reproduit ici) contrôlant l'âge, le niveau de vie, le niveau de diplôme, la situation d'activité et le sexe, montre qu'être dégroupé plutôt que figurer dans l'annuaire accroît de 6 points seulement la probabilité d'utiliser Internet tous les jours (au lieu de 17 points). Pour d'autres variables, le dégroupage conserve un effet propre important. Ainsi, toutes choses égales par ailleurs, être dégroupé accroît d'en général 20 points la probabilité de disposer d'un ordinateur chez soi.

Inversement, être sur liste rouge, et surtout être sans ligne fixe, réduit le degré d'équipement et de pratique. La population dans l'annuaire apparaît ainsi dans une situation intermédiaire du point de vue des TIC. C'est ce qui explique la remarquable proximité des moyennes observées dans l'annuaire et dans l'ensemble de la population.

En d'autres termes, une enquête TIC dans l'annuaire va aboutir à un biais très modéré, qui est la résultante de deux biais beaucoup plus prononcés, mais de sens opposés : si l'annuaire sous-estime nettement l'équipement et des pratiques des dégroupés, il sur-estime encore plus fortement ceux des MEM.

Ainsi, de manière peut-être inattendue, les évolutions de ces deux catégories ont conduit à une réduction du biais d'ensemble entre 2005 et 2008. La plate forme CAMME d'octobre 2005, réalisée sur une population de ménages dans l'annuaire, obtenait par exemple un taux moyen d'accès à Internet chez soi de 46,5% contre 41% dans la population totale. L'écart analogue n'est plus que de 1 point en 2008.

Tableau 4 : Quelques indicateurs TIC (niv. individu) selon l'équipement téléphonique et selon l'âge

		Ensemble	Dans l'annuaire	En liste rouge (chez FT)	Dégroupés	Sans ligne fixe
Ensemble de la population	Micro ordinateur à domicile	74	74	68	93	37
	Accès internet à domicile	68	69	64	91	15
	Envoi d'e-mails (depuis 3 mois)	57	56	51	73	32
	Cherche des informations pratiques sur Internet	19	19	16	27	7
	Recherche emploi sur Internet	38	36	35	50	20
	Accède à son compte bancaire	39	37	38	51	22
	Utilise Internet tous les jours	47	46	44	63	22
	Utilise un ordinateur tous les jours	54	52	48	69	36
Moins de 30 ans	Micro ordinateur à domicile	89	92	97	94	61
	Utilise Internet tous les jours	67	70	71	75	39
De 30 à 49 ans	Micro ordinateur à domicile	85	86	81	98	40
	Utilise Internet tous les jours	54	53	50	68	21
De 50 à 59 ans	Micro ordinateur à domicile	64	69	53	87	4
	Utilise Internet tous les jours	38	40	37	50	3
60 ans et +	Micro ordinateur à domicile	44	40	40	80	27
	Utilise Internet tous les jours	20	16	24	35	16
Diplôme primaire au plus	Micro ordinateur à domicile	62	64	53	87	25
	Utilise Internet tous les jours	31	33	24	47	10
Diplôme secondaire	Micro ordinateur à domicile	91	90	93	95	77
	Utilise Internet tous les jours	66	62	80	71	53
Diplôme supérieur	Micro ordinateur à domicile	91	92	86	99	65
	Utilise Internet tous les jours	73	72	69	81	54
Enquête 2005		Ensemble	PF CAMME			
	Micro ordinateur à domicile	56,5	56,5	nd	nd	39
	Accès internet à domicile	41	46,5	nd	nd	10

Source : Insee, enquête TIC 2008

Champ des individus : individus âgés de 16 à 74 ans, France métropolitaine.

Lecture : 74% des individus ont un ordinateur à domicile. 68% des individus vivant dans un ménage abonné à France Télécom et en liste rouge ont un ordinateur à domicile.

Naturellement, ce hasard heureux sur les résultats concernant l'ensemble de la population ne se retrouve pas nécessairement sur toutes les sous-populations qu'on est amené à examiner. Le résultat dépend du poids des dégroupés ou des MEM dans ces strates. Ainsi, en 2005, l'échantillon annuaire surestime de 8 points l'équipement en micro-ordinateur et de 15 points l'accès à Internet (14 points pour l'accès au haut débit) des familles monoparentales. Celles-ci sont aussi les moins pourvues en téléphone fixe : 25 % des familles mono-parentales ne disposent que d'un téléphone mobile.

De même, l'accès à internet des ménages modestes, plus souvent des mono-équipés mobiles (22 %), est surestimé par une enquête dans l'annuaire (+ 8 points).

Ces écarts ne sont évidemment pas redressables par recalage sur des marges (socio-démographiques). On est dans un cas typique de non-réponse non-ignorable : à caractéristiques fixées, un MEM aura presque "mécaniquement" beaucoup moins de chances d'être utilisateur d'Internet par exemple, puisque l'accès à Internet est moins aisé à partir d'un mobile. Symétriquement, appartenir à la population des dégroupés semble signaler, toutes choses égales par ailleurs, un intérêt particulier pour les TIC.

Remarque : échantillonner dans l'annuaire se révèle, on le voit, particulièrement problématique dans le cas des enquêtes sur les TIC, compte tenu du lien manifeste entre la situation de l'individu vis à vis de l'annuaire et son équipement et son comportement en matière informatique. Les enquêtes portant sur d'autres thèmes ne suscitent pas nécessairement autant d'interrogations. C'est le cas de l'enquête CAMME. Les réponses aux questions sur la conjoncture (opinion sur la situation économique générale, sur celle du ménage, intentions d'achat ou d'épargne, etc.) n'ont pas de raison particulière

d'être liées aux comportements TIC (à caractéristiques socio-démographiques du ménage données). C'est effectivement ce qu'on observe. Tirer l'enquête mensuelle de conjoncture dans l'annuaire ne semble occasionner aucun biais significatif (voir annexe).

### **3. Quel redressement pour les enquêtes TIC par téléphone ?**

#### **3.1 La méthode utilisée en 2007**

En 2007, l'enquête TIC a été effectuée auprès d'un échantillon de 7100 numéros de téléphone issus de l'annuaire. Pour redresser le biais identifié par l'enquête TIC de 2005, on a retenu la méthode décrite ci-dessous. Elle corrige l'échantillon de l'absence des monoéquipés en mobile. En revanche, compte tenu de l'information disponible dans l'enquête TIC 2005, elle doit faire l'hypothèse, erronée comme le montrent les résultats obtenus en 2008, que les ménages sur liste rouge et les ménages dégroupés ont le même comportement en matière de TIC que les ménages de l'annuaire.

Deux méthodes de redressement ont été envisagées. Toutes deux cherchent à utiliser l'information disponible dans l'enquête TIC d'octobre 2005 :

- la méthode de redressement par repondération opère sur les taux d'équipement ou de pratique. Par exemple, le taux de possession d'ordinateur calculé en 2007 (sur la seule population figurant dans l'annuaire) est redressé de l'écart observé en 2005, sur cette variable, entre la population de l'annuaire et la population des mono-équipés de mobile. Pour une variable prise séparément, cette méthode est la plus simple, et permet de calculer la variance de sondage. En revanche, si on doit produire de nombreuses estimations de taux d'équipement ou de pratiques, qui peuvent être liés entre eux, et pour différents groupes de la population (par âge, par diplôme, etc.), il est pratiquement impossible d'assurer la cohérence de tous les redressements.

- la méthode de redressement par imputation, elle, assure cette cohérence. Elle consiste à produire un échantillon complémentaire de personnes MEM auxquelles on impute les réponses aux questions posées en 2007. L'enquête CVS de 2007, réalisée au premier trimestre 2007, auprès d'un important échantillon (17 000 individus), constituait une source adaptée où aller chercher cet échantillon. Elle comporte en effet des variables qui renseignent sur la possession de ligne fixe et de téléphone(s) portable(s) par les ménages. CVS 2007 a ainsi été l'enquête "donneuse" d'un échantillon d'individus de 15 ans et plus (vivant dans des ménages MEM) accompagnés de leurs caractéristiques socio-démographiques. Deux types de variables ont été imputés :

- les variables de TIC 2007 qui existaient déjà dans l'enquête PCV d'octobre 2005. On a effectué des régressions logistiques sur la population des monoéquipés mobiles de 2005, en prenant comme régresseurs les variables socio-démographiques de ces individus. Les équations estimées ont été ensuite utilisées pour prédire les réponses des individus « imputés » de 2007. Un coefficient d'évolution a ensuite été appliqué à ces réponses, afin que les évolutions globales 2005-2007 soient les mêmes sur la population des MEM que sur la population disposant d'un téléphone fixe.
- les variables de TIC 2007 qui n'existaient pas dans l'enquête PCV d'octobre 2005 : dans le cas de ces variables, il a fallu supposer que les ménages MEM avaient le même comportement que les ménages disposant d'une ligne fixe, et présentant les mêmes caractéristiques socio-démographiques. Les réponses des individus imputés ont été prédites à partir de celles des répondants de TIC 2007. Ici l'approche par régression revient à un recalage sur marges.

Après les imputations, l'échantillon global (répondants + imputés) a été calé sur les marges habituelles et sur une variable indiquant l'origine de l'individu (échantillon téléphonique ou population imputée).

#### **3.2 L'ampleur du redressement**

Comme on le montrent les résultats des expériences de 2005 et de 2008 présentés plus haut, les écarts entre les ménages dans l'annuaire et ceux n'y figurant pas sont limités pour l'ensemble de la population, plus forts sur des sous-catégorie, et recouvrent en fait deux biais plus forts mais de sens contraire et qui tendent à se compenser.

Dans ces conditions, le redressement reste modéré, au niveau de l'ensemble de la population. Il

affecte assez peu la plupart des variables, notamment celles associées aux usages d'Internet (voir graphiques de la section 3.3). Il a cependant un impact notable sur d'autres, comme le taux d'accès chez soi à un ordinateur ou à Internet (qui se trouvent par ailleurs être des indicateurs particulièrement suivis dans les utilisateurs des données sur les TIC).

Tableau 5 : accès à Internet et équipement en micro-ordinateur au niveau individuel

population	Accès à Internet		Équipement en micro-ordinateur	
	2007 redressé	2007 observé	2007 redressé	2007 observé
12-17	81	81	92	93
18-24	73	81	84	89
25-39	66	75	80	85
40-59	62	66	73	76
60-69	36	37	46	47
70 et +	9	9	14	14
Ens	53	56	63	65

Source : Insee, enquête TIC 2007

Champ : individus vivant dans les ménages ordinaires

Lecture : 66% des individus de 25 à 39 ans vivent dans un ménage ayant accès à Internet.

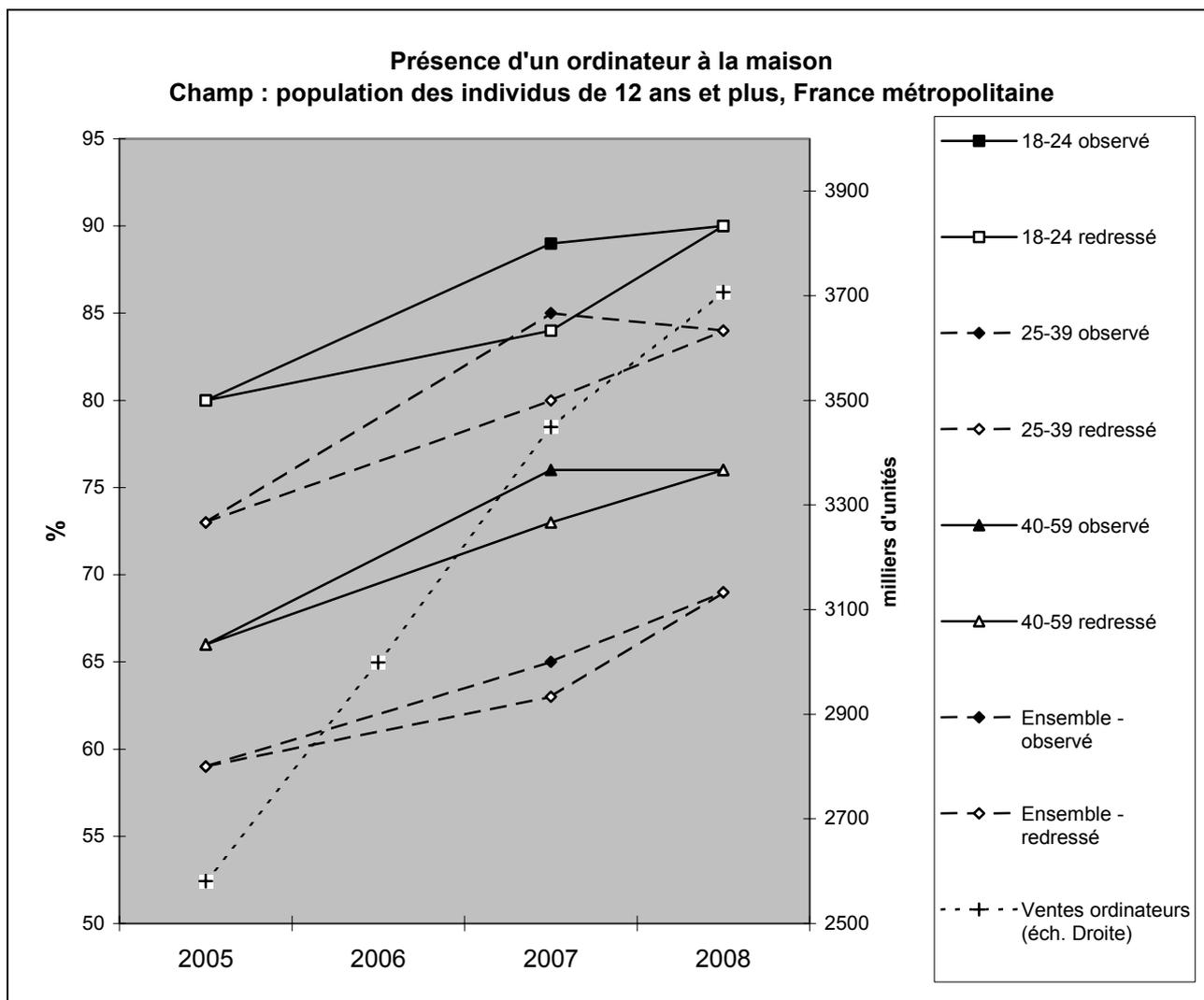
Le redressement consiste en 2007 à "réinjecter" les MEM manquants dans l'échantillon. Ceux-ci se distinguant par un équipement (ordinateur ou accès internet) plus faible et une pratique des TIC moins développée, le redressement conduit, en général, à diminuer les taux observés dans l'échantillon tiré de l'annuaire. L'effet est le plus fort sur les tranches d'âge où les monoéquipés mobiles sont relativement plus nombreux (essentiellement les ménages les plus jeunes).

Remarque : la mesure de l'équipement est effectuée au niveau du ménage. Les individus les plus jeunes (12-17 ans et une partie de 18-24) vivent généralement chez leurs parents. Ceux-ci, le plus souvent proches ou au-dessus de 40 ans, ont plus de chance de disposer d'un téléphone fixe. Le taux d'équipement des plus jeunes sera donc peu affecté par le redressement. Ce sont les 18-40 ans qui ont le plus de chance d'être des MEM et donc d'être les plus affectés par le redressement. Par exemple, on observe cinq points d'écart pour les 18-40 ans en ce qui concerne l'équipement en micro-ordinateur, et seulement un point d'écart chez les plus de 60 ans. Pour l'accès à Internet, le redressement conduit à un écart de 8 à 9 points chez les 18-40 ans, contre un point pour les plus de 60 ans.

### 3.3 Comment évaluer la pertinence du redressement ?

#### 3.3.1 Comparaison avec des sources internes

Nous disposons de sources de données comparables à celles de 2007 : la plupart des indicateurs étaient déjà disponibles en 2005, et l'ont été à nouveau en 2008. Nous pouvons donc observer deux types d'évolution : une avec les chiffres de 2007 redressés, et l'autre avec le chiffre de 2007 observé (voir exemple ci-dessous et graphiques en annexe).



Source : Insee, enquêtes PCV d'octobre 2005, TIC 2007 et TIC 2008.

A première vue, le redressement tendrait à diminuer le pourcentage global d'individus disposant d'un ordinateur à domicile en 2007 : le chiffre observé s'inscrit dans une évolution parfaitement linéaire entre 2005 et 2008, tandis que le chiffre redressé est légèrement plus bas. Mais si on étudie ces évolutions par tranche d'âge, on s'aperçoit que l'effet global n'est pas le même partout :

- chez les 18-24 ans, le redressement fait grandement diminuer le pourcentage des possesseurs d'ordinateurs en 2007, mais l'évolution devient alors plausible : si on se contente du chiffre observé, on conclue à une stagnation de ce chiffre entre 2007 et 2008, qui paraît peu probable au vu de l'évolution des ventes d'ordinateurs. Cette stagnation disparaît après redressement.
- chez les 25-39 ans et les 40-59 ans, le redressement « linéarise » l'évolution entre 2005 et 2008, et supprime la légère baisse du pourcentage que l'on aurait observée en l'absence de redressement, et qui paraissait peu crédible.

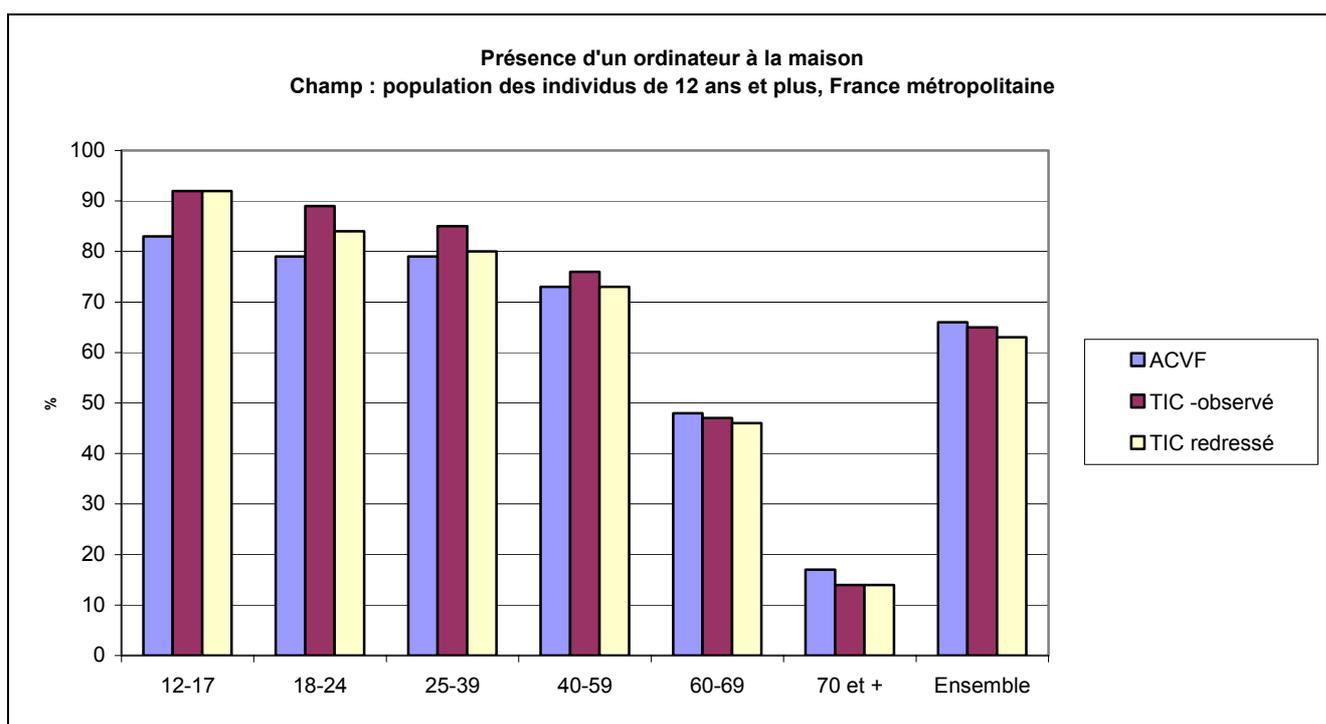
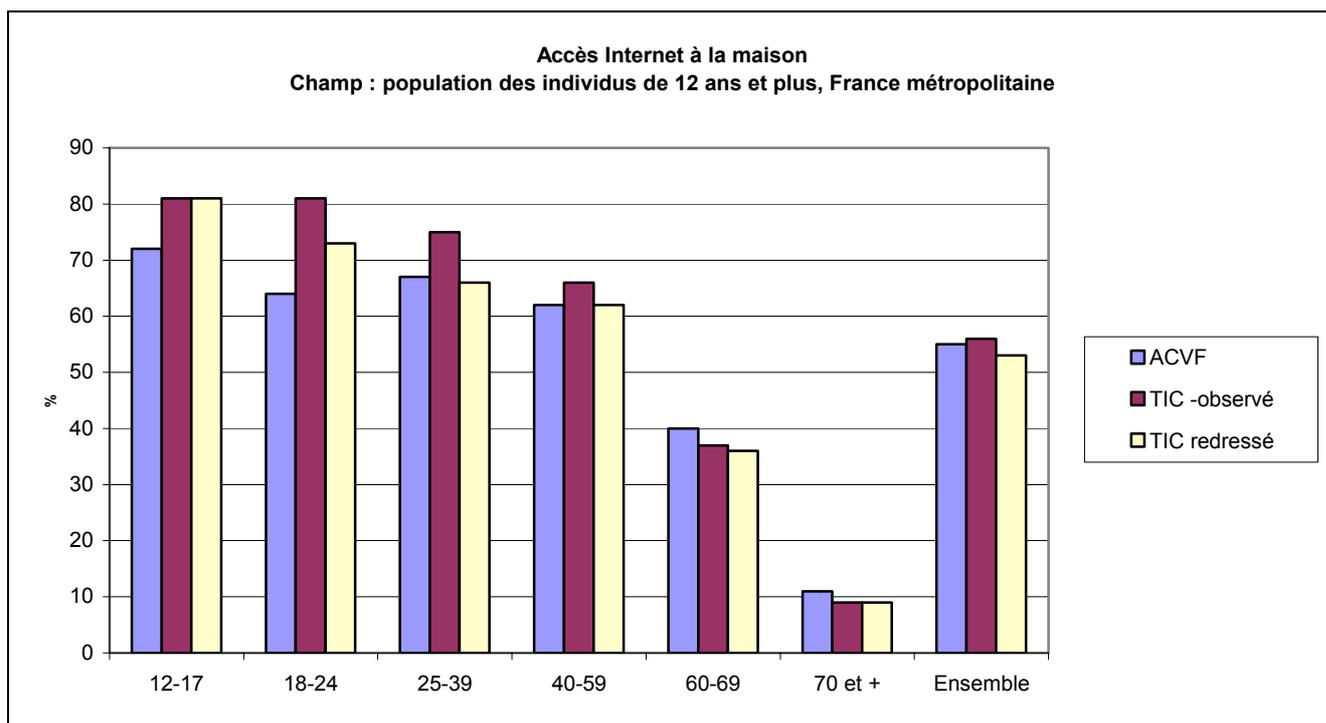
L'étude menée sur d'autres variables comme l'accès internet, l'envoi d'e-mails, la consultation en ligne du compte bancaire et l'utilisation d'internet pour obtenir des informations administratives ou des formulaires, aboutit aux mêmes conclusions.

En outre, on observe à plusieurs reprises, une tendance du redressement à "trop" diminuer les taux calculés sur l'échantillon annuaire. L'interprétation la plus évidente de ce phénomène consiste à noter que le redressement n'a pu tenir compte que des mono-équipés mobile. En particulier, il ne distingue pas les dégroupés, et les aligne sur les individus de l'annuaire. Or, ceux-ci ont des taux nettement plus élevés, et qui évoluent plus vite que ceux des ménages et individus de l'annuaire. Autrement dit le redressement corrige la sur-estimation liée à l'absence des monoéquipés de mobiles dans l'annuaire, mais pas la sous-estimation liée à l'absence des dégroupés (l'absence des listes rouge n'ayant pas d'incidence particulière, comme on l'a vu).

Au total, le redressement semble « améliorer » les résultats de 2007, dans la mesure où il tend à rendre les évolutions plus linéaires. Mais, bien entendu, rien ne garantit que les vraies évolutions 2005-2008 soient rectilignes. Il convient donc d'examiner une autre méthode de contrôle des résultats.

### 3.3.2 Comparaison avec des sources externes

Les enquêtes "Aspirations et conditions de vie des Français" (ACVF), réalisées chaque année par le CREDOC, abordent le domaine des TIC. Elles collectent en particulier des données directement comparables avec les principaux indicateurs recueillis dans les TIC. Elles fournissent ainsi de très utiles éléments de comparaison.



Sources : Insee, enquête TIC 2007  
Credoc, enquête "Aspirations et conditions de vie des Français" 2007

Au niveau de l'ensemble de la population, le redressement semble éloigner légèrement les résultats de l'enquête TIC de ceux de l'ACVF (les deux chiffres restant assez proches). Mais là encore, ce n'est pas le même effet observé si on regarde les tranches d'âge :

- chez les 12-17 ans ou les plus de 60 ans, le redressement n'apporte rien : en effet, ces deux populations comportent peu de ménages MEM. Ainsi, les enquêtes ACVF fournissent des pourcentages au-dessus de ceux de TIC pour les plus vieux, et au-dessous pour les plus jeunes, avec ou sans redressement.
- chez les 18-59 ans, le redressement va dans le sens attendu en baissant les chiffres de TIC, les rapprochant de ceux d'ACVF.

Ainsi, le redressement a un effet améliorant (ou nul) sur toutes les tranches d'âges, même s'il éloigne les taux des deux enquêtes sur l'ensemble de la population.

Ces deux méthodes d'étude de l'efficacité conduisent à la même conclusion : le redressement en 2007 améliore les résultats issus d'un échantillon tiré dans le seul annuaire. Il ne résout cependant pas toutes les difficultés, notamment, faute d'information suffisante en 2005, le biais induit par les dégroupés. Dans ces conditions, les résultats de l'enquête TIC 2007 sont affectés d'une marge d'incertitude plus élevée que ceux des autres années. Il importe en particulier d'être prudent dans l'analyse des inflexions sur 2005-2007 et sur 2007-2008.

## Bibliographie

Bigot R., Croutte P. (2008), *La diffusion des technologies de l'information et de la communication dans la société française*, Rapport pour le CGTI et l'Arcep, Credoc, novembre 2008.

Roy G., Vanheuverzwyn A.(2003), « Le téléphone mobile dans les enquêtes par sondage », in *Actes des journées de méthodologie statistique des 4 et 5 décembre 2000*, Insee Méthodes n° 102, T. 2, pp 109-122, Insee.

## Annexe : Résultats de l'opération méthodologique d'avril 2008 pour l'enquête CAMME

L'échantillon hors annuaire était interrogé d'une part sur le questionnaire TIC, d'autre part sur le questionnaire CAMME d'avril. L'interrogation a eu lieu exactement durant la période d'enquête CAMME (3 premières semaines du mois). On dispose donc des réponses d'individus en principe absents des échantillons enquêtés par CAMME. Il est ainsi possible de comparer les résultats de l'enquête de conjoncture sur les deux échantillons.

On a comparé sur les deux échantillons les 11 indicateurs de conjoncture présentés mensuellement par la division des comptes trimestriels dans les *Informations rapides*, et l'indicateur résumé ("moral des ménages"). On a en outre retenu les deux indicateurs d'opinion sur la situation économique générale (passée et future).

On rappelle que ces indicateurs (hormis l'indicateur résumé) prennent 3 valeurs (-1, 0, 1), la valeur négative correspondant à l'opinion "pessimiste". Ils sont déduits de questions en général à 4 ou 5 modalités<sup>12</sup>.

Les écarts dans les réponses des différents groupes de ménages (annuaire/liste rouge/dégroupés/sans fixe) sont de bien moindre ampleur que dans le cas de l'enquête TIC. A la différence du cas TIC, ils changent de signe selon l'indicateur de façon assez peu interprétable, même si les dégroupés apparaissent un peu plus optimistes que les autres.

Surtout, ils s'expliquent dans pratiquement tous les cas par des différences de structure socio-démographique : l'analyse des effets attribuables, toutes choses égales par ailleurs (i.e. ici à âge, sexe, niveau de vie, niveau d'éducation, statut d'activité fixés), à la situation du ménage vis à vis de l'annuaire montre très nettement l'absence de lien entre les réponses des ménages à l'enquête de conjoncture et le seul fait d'être sur liste, dégroupé, ou sans fixe.

Tableau 2 : valeurs des indicateurs CAMME selon l'équipement téléphonique

		Ensemble	Annuaire	Sur liste		Dégroupé		Sans fixe	
				niveau	effet spécif. (+)	niveau	effet spécif. (+)	niveau	effet spécif. (+)
situation économique générale	passée	-72	-74	-77	(*)	-66		-66	
	future	-36	-39	-35		-32		-30	
Situation financière personnelle	actuelle	8	13	11		9		-16	
	passée	-28	-30	-28		-17		-32	
	future	-10	-15	-11		0		2	(***)
capacité à épargner		-11	-14	-7		2		-18	
opportunité d'acheter		-31	-30	-28		-32		-38	
opportunité d'épargner		26	23	27		31		30	
Niveau de vie des français	passé	-70	-72	-73		-68		-63	
	futur	-41	-43	-42		-35		-40	
Chomage futur		-34	-33	-33		-34		-35	
Evolution des prix	passée	37	39	35	(***)	31		41	(*)
	future	-60	-55	-70		-61		-67	
<b>Indicateur résumé</b>		<b>-36</b>	<b>-38</b>	<b>-36</b>		<b>-30</b>		<b>-34</b>	

source : Enquête TIC avril 2008, enquête CAMME avril 2008

Champ : ménages ordinaires, France métropolitaine

Note : (+) présence d'un effet spécifique déterminée par régression logistique sur des caractéristiques du ménage

(\*) effet significatif au seuil de 20%

(\*\*) effet significatif au seuil de 10%

(\*\*\*) effet significatif au seuil de 5%

<sup>12</sup> Par exemple, la question sur le chômage ("Pensez-vous que, dans les douze prochains mois, le nombre de chômeurs va : fortement augmenter/un peu augmenter/rester stationnaire/un peu diminuer/fortement diminuer/ ne sait pas) donne lieu à un indicateur qui vaut 1 si le répondant parle de diminution, -1 si il parle d'augmentation, 0 dans les autres cas.