

Redresser la non-réponse totale dans les enquêtes auprès des entreprises : les pièges à éviter

Illustration avec l'enquête innovation.

Philippe BRION^(), Nathalie CARON^(**) et Pascale PIETRI-BESSY^(*)*

() Insee, Direction des statistiques d'entreprises, (**) Insee, Unité méthodes statistiques*

1. Introduction

Corriger la non-réponse dans une enquête n'est jamais chose simple et évidente. Des techniques de correction existent, en particulier la repondération et l'imputation¹, mais elles nécessitent pour être appliquées correctement de tenir compte du contexte dans lequel on travaille. En effet, leur utilisation de façon mécanique, un peu comme des « boîtes noires », permettrait effectivement de se sortir de n'importe quelle situation, mais avec le risque d'introduire du biais.

Les enquêtes menées par le système statistique public français auprès des entreprises sont marquées par plusieurs spécificités.

Premièrement, elles sont réalisées par courrier. Ceci veut dire qu'on ne maîtrise pas le processus temporel de retour des questionnaires d'une part, et qu'on n'a pas en général de contact direct avec les entreprises échantillonnées. Les entreprises non répondantes sont bien sûr « relancées », en général par courrier, mais il arrive un moment où il est nécessaire de décider que le processus de relance est terminé, et qu'on arrête le processus de collecte. Notons que les entreprises « importantes » (en termes de chiffre d'affaires et/ou d'effectif) sont en général relancées en face-à-face par le réseau des enquêteurs entreprises.

Le traitement des non-réponses totales pose alors un problème : comment savoir si un questionnaire non retourné correspond à un véritable refus de répondre, ou au fait que l'unité enquêtée ne se sent pas concernée car ne faisant plus partie du champ de l'enquête, par exemple si elle a changé d'activité, ou encore si elle a cessé son activité ? On ne peut donc pas assimiler "non-retour" et "non-réponse".

Deuxièmement, se pose la question de la base de sondage et de son utilisation potentielle pour les redressements, par exemple via une méthode de calage. Les méthodes d'échantillonnage utilisées en France pour les enquêtes statistiques menées auprès des entreprises s'appuient sur le répertoire Sirene. Le fait que celui-ci soit un répertoire inter-administratif place la statistique publique française dans une situation relativement confortable : pour démarrer son activité, toute entreprise doit disposer d'un n° Siren, et les questions de sous-couverture de la base de sondage ne se posent pas. Par contre, comme dans tout répertoire, il existe un problème d'ancienneté des informations qui conduit à l'existence de faux actifs (unités cessées mais non enregistrées comme

¹ Pour plus de détails, se reporter par exemple aux références bibliographiques [2] et [4].

telles), ou à de mauvais classements d'unités² (en particulier en fonction de l'activité principale de l'entreprise).

Un certain nombre de pièges jalonnent alors le parcours du chargé d'enquête, qui, disposant à la fois d'un répertoire fournissant des distributions et de méthodes de redressement « reconnues », doit procéder au travail de production de résultats.

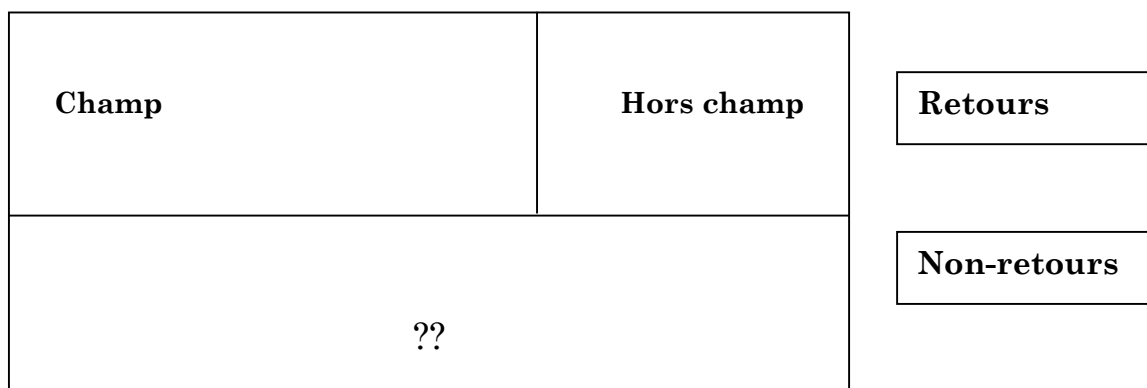
L'objet de ce papier est de présenter les difficultés potentielles existant dans le cadre de ce processus ainsi que les principes à adopter pour les traiter (partie II), et d'illustrer ceci sur un cas concret, celui de l'enquête sur l'innovation dans les services (partie III).

2. Les problèmes

2.1. Non-réponse et non-retour : deux notions différentes

Lorsqu'un statisticien réalise une enquête, il se trouve dans la situation suivante : pour certaines des unités échantillonnées (que l'on appellera par la suite les « retours ») il reçoit de l'information, et pour les autres (que l'on appellera par la suite les « non-retours ») il ne reçoit aucun questionnaire en retour. Cette catégorie des « non-retours » contient en fait deux groupes d'unités : celles qui font partie du champ de l'enquête pour lesquelles on aurait dû recevoir un questionnaire (c'est-à-dire les non-réponses) et celles qui ne font pas partie du champ, en particulier les cessations.

Figure 1 : Répartition des questionnaires après réalisation de l'enquête



On utilisera ici la définition suivante : une unité non-répondante est une unité échantillonnée qui appartient au champ de l'enquête et pour laquelle on n'a aucune réponse au questionnaire. Cet ensemble d'unités non répondantes ne correspond pas exactement à celui des « non-retours ». Autrement dit, les non-réponses sont incluses dans les non-retours mais la réciproque n'est pas vraie.

Cette dichotomie entre non-réponse et non-retour s'explique principalement par les deux raisons suivantes :

- il existe toujours un décalage temporel entre les informations de la base de sondage et celles recueillies lors de la collecte,
- les informations disponibles dans la base de sondage ne permettent pas toujours de cerner précisément la population ciblée par l'enquête.

² qui peuvent s'avérer pénalisants quand on réalise une enquête sur un champ donné (défini en termes de secteur économique, ou en termes de taille d'entreprises).

Pour illustrer ces situations, prenons deux exemples, le premier concernant les enquêtes entreprises et le second les enquêtes ménages.

① La base de sondage des enquêtes entreprises est constituée à partir du répertoire Sirene ; Elle est susceptible de comporter certaines entreprises considérées dans le champ de l'enquête au moment du tirage de l'échantillon qui s'avèrent au moment de sa réalisation hors champ³. Ainsi, un certain nombre de « fausses actives » qui existent à tort dans la base de sondage se sont retrouvées sélectionnées dans l'échantillon. De même, certaines entreprises ont pu changer de secteur depuis la dernière mise à jour de Sirene, ce qui peut les faire sortir du champ d'une enquête. Une entreprise donnée peut alors ne pas renvoyer son questionnaire, soit parce qu'elle n'existe plus, soit parce qu'elle n'est plus concernée.

② Le champ des enquêtes ménages correspond en général à l'ensemble des personnes qui vivent dans des résidences principales au moment de la réalisation de l'enquête. Or, la base de sondage principale de ces enquêtes est une base de logements constituée une fois pour toutes après chaque recensement. Comme le statut des logements de cette base de sondage ne peut pas être mis à jour, certains des logements sélectionnés pour une enquête s'avèrent par conséquent être des résidences secondaires, vacantes ou occasionnelles lors de la réalisation de l'enquête et constituent donc des logements hors champ.

La réalité est encore plus complexe car outre la distinction entre non-réponse et non-retour, il est aussi nécessaire d'« isoler » parmi les retours les réponses retenues pour participer aux résultats de l'enquête. En effet, ce n'est pas parce que l'on dispose d'un questionnaire en retour que les informations contenues dans ce questionnaire sont automatiquement retenues pour l'exploitation de l'enquête. En particulier, certains questionnaires remplis ne sont pas tous « exploitables » car ils ne contiennent pas suffisamment d'informations fiables, d'autres correspondent à des unités qui s'avèrent hors champ au vu de leurs réponses à l'enquête (changement d'APE pour une enquête sectorielle, ou franchissement de seuil, par exemple). Autrement dit, la catégorie des unités pour lesquelles on possède un questionnaire en retour n'est pas homogène.

2.2. Une étape importante : la mobilisation de sources externes

Comme indiqué précédemment, les non-retours posent problème, et on ne peut pas utiliser le répertoire pour connaître l'état des unités correspondantes si on procède au redressement juste après avoir réalisé l'enquête (et donc peu de temps après avoir tiré l'échantillon). Cependant, comme on le verra dans la partie III, le répertoire peut apporter des informations intéressantes si suffisamment de temps s'est écoulé entre la réalisation de l'enquête et le travail de redressement.

Une solution pour affiner la connaissance relative aux non-retours peut être de mener une investigation complémentaire, plus légère que l'enquête : c'est le cas des enquêtes d'amélioration du répertoire (EAR) menées par les gestionnaires du répertoire, à la demande des responsables d'enquêtes, en particulier pour les enquêtes annuelles d'entreprise (EAE), qui consistent à contacter l'unité afin de connaître son caractère actif ou non, ou à utiliser différentes informations ayant pu être connues par ces gestionnaires (par exemple postérieurement au tirage de l'échantillon). Notons que les EAR elles-mêmes comprennent de la non-réponse ...

Par ailleurs, des sources complémentaires et beaucoup plus exhaustives existent et sont utilisées depuis quelques années, en particulier pour les EAE : les sources fiscales. Ces sources permettent d'affiner la connaissance des entreprises, au moins sur leur caractère actif ou non, et de réduire par conséquent la taille de la catégorie des non-retours pour laquelle des incertitudes sur le statut persistent. La part « d'inconnu » qui reste devra ensuite être traitée lors du redressement.

³ Même si la période sur laquelle l'enquête est antérieure à sa date de réalisation (par exemple les EAE réalisées l'année n+1 portent sur l'année d'exercice n).

D'une part, l'Insee a accès au fichier des déclarations de TVA des entreprises, et on dispose ainsi d'informations sur l'activité des entreprises présentes dans ce fichier qui peut être constitué, relativement à une année n , en juin $n+1$.

D'autre part, il existe une liste des entreprises, sous forme de fichier, pour lesquelles des déclarations sur l'exercice n sont attendues par la direction générale des impôts. Cette liste est disponible en septembre $n+1$. Ce fichier est relatif aux déclarations annuelles des entreprises (soumises au régime des bénéfices industriels et commerciaux, ou à celui des bénéfices non commerciaux) et également à d'autres impôts (TVA, mais aussi d'autres taxes comme la taxe d'apprentissage). Il contient une variable sur la date de cessation potentielle de l'entreprise.

2.3. Les pièges à éviter lors du redressement de l'enquête

Les responsables d'enquête sont souvent tentés de « retrouver », à partir de l'échantillon de l'enquête, des distributions ou « masses » connues. Pour cela ils ont parfois tendance à partir du fichier qu'ils ont obtenu, après traitement des questionnaires reçus, et à y appliquer des procédures de repondération. Ceci veut dire qu'ils ne gardent que les questionnaires reçus, et jugés suffisamment cohérents, pour procéder à ce travail de repondération (éventuellement opéré au travers d'un calage sur marge), en s'appuyant sur une distribution connue en général au travers du répertoire d'entreprises.

Cette manière de procéder n'est évidemment pas correcte puisque la « cohérence » entre l'échantillon qu'on veut repondérer, composé exclusivement d'unités vivantes et exploitables, et une population de référence qui est a priori susceptible de contenir des unités cessées n'est pas assurée. Il est donc nécessaire de ne pas se limiter aux seules unités de l'échantillon jugées utilisables, mais de repartir d'un ensemble d'unités plus étendu. Le traitement des non-réponses totales (en fait des non-retours) nécessite alors de définir de manière précise des sous-populations sur lesquelles procéder à la repondération ou au calage, présentées au paragraphe suivant.

Par ailleurs, en cas de redressement par repondération utilisant une source sur laquelle on se cale, le statisticien doit veiller à assurer une cohérence des variables de calage entre le fichier de l'enquête et la source externe. Autrement dit, ce doit être les mêmes variables, c'est-à-dire les mêmes concepts recueillis à une même date, qui sont utilisées à la fois pour constituer les marges et dans le fichier d'enquête « à caler ». Ceci signifie que, dans notre cas, nous utiliserons comme données de calage des valeurs existant dans une source de référence comme le répertoire qui est une source exhaustive, et non les valeurs de variables observées à l'enquête (qui par essence même ne sont connues que sur l'échantillon)⁴.

2.4. Une correction de la non-réponse qui doit s'adapter à la réalité

Comme nous l'avons déjà signalé dans la partie II.2, la mobilisation de sources externes est une étape très importante du processus de correction de la non-réponse car elle permet d'apporter de l'information sur le statut (dans le champ de l'enquête ou non) des unités de la catégorie « non-retours ». Cependant, dans la pratique pour les enquêtes entreprises réalisées par voie postale, il est quasi impossible d'obtenir de l'information fiable sur chacune des unités de cette catégorie ; il subsistera donc toujours des unités pour lesquelles il sera impossible de définir le statut.

Par conséquent, les unités échantillonnées peuvent être réparties dans les quatre catégories suivantes :

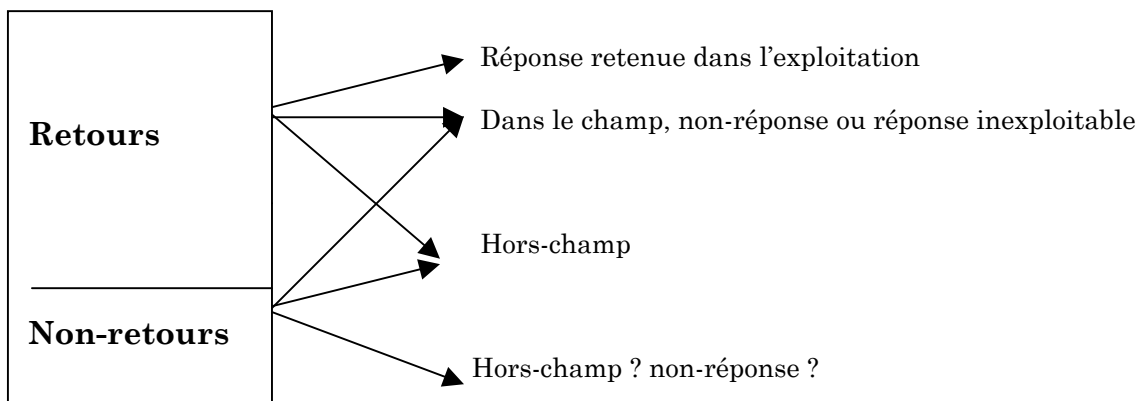
1. **les unités répondantes** (appartenant au champ de l'enquête),
2. **les unités repérées comme non-répondantes** : les unités dont le questionnaire est jugé inexploitable et les unités identifiées comme non-répondantes parmi les non-retours par une source externe.

⁴ Pour plus de détails, se reporter à [3].

3. **les unités repérées comme hors champ** : les unités classées comme hors champ pour changement d'activité ou pour cessation d'après leurs réponses au questionnaire⁵ ainsi que celles identifiées comme cessées parmi les non-retours par une source externe.
4. **les autres unités de la catégorie « non-retours » pour lesquelles on ne sait rien, qui sont soit hors champ soit non répondantes.**

Il est important de noter que les entreprises qui ont cessé leur activité constituent un cas important de la catégorie 3. De plus, les catégories 2 et 3 contiennent aussi bien des unités pour lesquelles on a eu un questionnaire en retour que des unités pour lesquelles on n'en a pas eu.

Figure 2 : Répartition des questionnaires en quatre catégories pour le traitement de la non-réponse



Cette répartition de l'échantillon en quatre catégories est primordiale avant de se lancer dans la correction de la non-réponse totale proprement dite, quelle que soit la méthode retenue, par repondération ou par imputation.

Ainsi, par exemple, si on opte pour une méthode de repondération, la modification des pondérations doit s'effectuer en deux étapes successives :

- la première étape va consister à augmenter le poids des unités des trois premières catégories pour compenser les unités de la dernière catégorie celle des « non-retours sans information auxiliaire ». Elle traduit le fait que ces dernières unités sont susceptibles de relever des catégories 1 à 3.
- la seconde visera à augmenter le poids des unités de la première catégorie pour compenser ceux de la seconde. Elle opère totalement sur les unités appartenant au champ.

Ces deux étapes peuvent être réalisées par une méthode classique de correction de la non-réponse par repondération comme la méthode des groupes de réponse homogène ou encore par calage. **La première étape est primordiale** car elle vise à « attribuer » une partie des unités de la quatrième catégorie dans le champ de l'enquête et l'autre partie hors du champ de l'enquête par la modification des pondérations. Si on oublie cette étape et que la correction de poids adoptée consiste uniquement à augmenter le poids des unités de la première catégorie pour compenser ceux de la seconde et de la dernière catégories, le nombre d'unités estimées dans le champ de l'enquête sera sur-estimé. En effet, par construction, cette solution revient à considérer que toutes les unités de la catégorie des « non-retours sans information auxiliaire » sont dans le champ de l'enquête alors qu'en réalité ce n'est pas le cas. La principale difficulté relative à la première étape consiste à « travailler » sur le bon fichier qui correspond ici à l'ensemble des unités du fichier en n'oubliant pas en particulier les unités hors du champ de l'enquête. Quant à la seconde étape, elle est plus traditionnelle et correspond exactement à ce qui est fait pour corriger la non-réponse

⁵ Ou identifiées comme telles lors de la collecte par les gestionnaires.

dans les enquêtes ménages puisque l'on arrive avec les informations des enquêteurs à attribuer un statut (dans le champ ou non) à tous les logements sélectionnés. La catégorie 4 est donc vide dans ce cas et par conséquent la première étape n'existe pas.

De façon similaire, si on choisit une technique d'imputation plutôt que celle de repondération, on se heurte au même problème. Pour les unités pour lesquelles on ne sait rien, il est nécessaire de leur imputer un statut. Autrement dit, si on réalise par exemple une imputation par hot-deck, le « donneur choisi » pour une unité de la dernière catégorie peut être soit une unité répondante dans le champ de l'enquête soit une unité hors champ⁶. En revanche, pour les unités repérées non-répondantes (catégorie 2), le donneur choisi sera forcément une unité répondante dans le champ de l'enquête.

Ajoutons que ce processus de correction ne changera en rien l'exploitation de l'enquête qui ne portera naturellement que sur les unités de la première catégorie, c'est-à-dire les unités répondantes dans le champ de l'enquête.

3. Application à l'enquête Innovation - quelques enseignements

3.1. Présentation de l'enquête innovation

L'enquête « **innovation générale** » dite aussi « innovation française dans les services » a été conduite en 2001/2002 auprès d'environ 2 400 entreprises. Cette enquête avait pour vocation de mieux cerner ce que les entreprises de services considéraient comme une innovation (voir [1]). Le champ sectoriel recouvre les secteurs des « activités juridiques, comptables et de conseil de gestion », des « services opérationnels (nettoyage, assainissement) », de la publicité, de l'hôtellerie, de la restauration, des « activités récréatives et culturelles » (audiovisuel et agences de presse).

Le plan de sondage de l'enquête est stratifié à un degré, la stratification croisant les critères « secteur d'activité » et « taille de l'entreprise ». Il s'est appuyé sur le plan de sondage de l'EAE, puisqu'on a sous-échantillonné à l'intérieur de l'échantillon de cette dernière enquête de façon quasi-homothétique (en en tirant une sur treize, à de rares exceptions près où on a appliqué des taux plus forts). Le plan de sondage de l'EAE utilisant des taux différenciés selon la taille des entreprises (avec un seuil d'exhaustivité situé entre 20 et 50 salariés selon le secteur d'activité), les taux de sondage sont plus importants dans les strates correspondant aux grandes entreprises. Au final, la pondération d'une unité de l'enquête innovation est calculée comme le produit de la pondération de cette même unité dans l'EAE par l'inverse du taux de sous-échantillonnage de l'EAE (en général 13).

Avant de procéder à la correction de la non-réponse proprement dite, il est nécessaire, comme nous l'avons vu précédemment, de classer les entreprises sélectionnées dans les différentes catégories présentées ci-dessus. Pour cela, nous disposons à l'issue de la collecte de deux fichiers : le fichier contenant l'ensemble des unités sélectionnées dans l'échantillon et celui dit « des questionnaires ». Le nombre d'enregistrements du fichier de tirage est égal au nombre d'enregistrements du fichier « des questionnaires ». Ce dernier fichier contient en particulier la variable de gestion de la collecte habituelle des enquêtes auprès des entreprises (dite code gestion d'état de la collecte) qui va nous permettre de réaliser le classement en quatre catégories.

Cependant, ce code n'est pas directement adapté. En effet, il sert comme son nom l'indique pour la gestion de la collecte et évolue au fur et à mesure de l'état d'avancement de la collecte. Dans le fichier, nous disposons de la variable en l'état de renseignement dans lequel elle se trouve lorsque la collecte est arrêtée. Au lancement de l'enquête, tous les questionnaires sont classés dans la modalité « questionnaires dits envoyés ». Au fur et à mesure du déroulement de la collecte, les questionnaires sont qualifiés par une autre modalité (par exemple, « cessations d'entreprises » ou bien « questionnaires dits reçus »). En fin de collecte, tous les questionnaires de la modalité « dits

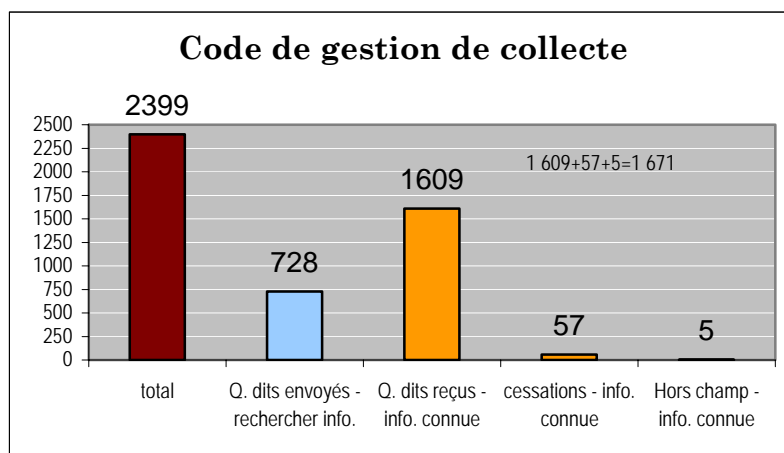
⁶ C'est ce que l'on appelle dans la chaîne de traitement des Enquêtes annuelles d'entreprise « l'imputation sans précision de champ »

envoyés » pour lesquels aucune information n'est revenue en retour restent dans cette modalité, c'est un peu comme une bouteille qui se serait vidée dans d'autres bouteilles, mais dans laquelle il reste toujours un fond.

Il serait préférable de créer une seconde variable plus opérationnelle, d'état en fin de collecte, où les questionnaires de la modalité du code de gestion « questionnaires dits envoyés » seraient classés dans une modalité « questionnaires en non retour », les autres modalités de cette variable pouvant être égales aux autres modalités de la variable de gestion qui sont moins ambiguës.

La variable de gestion de la collecte n'est donc pas suffisante car elle ne permet pas d'isoler rapidement les bons questionnaires des autres. Et il ne faudrait surtout pas faire une sélection des enregistrements correspondant aux questionnaires pour lesquelles on dispose d'une information en retour de collecte (soit le questionnaire est rempli, soit l'entreprise est cessée, soit l'entreprise est hors-champ (changement d'activité) et laisser tomber les autres. Sinon ce serait supposer que la population des questionnaires en non retour est en tout point identique à celle des questionnaires en retour. Dans notre exemple (voir figure 3), on a 1 671 questionnaires retournés et 728 questionnaires envoyés pour lesquels on n'a rien reçu, ou pas suffisamment d'information. Il va alors s'agir de définir les 4 catégories identifiées ci-dessus en recourant à d'autres sources pour lever certaines ambiguïtés.

Figure 3 : Répartition des questionnaires selon les modalités du code de gestion de collecte



3.2. Mobilisation de sources externes

L'existence de ces 728 entreprises pour lesquelles nous n'avons pas d'information en retour de collecte nous a conduit à mobiliser des données issues d'une source externe, en l'occurrence le répertoire SIRENE (la mobilisation de sources fiscales, présentée au II.2, n'a pas été utilisée ici). Ces données sur la démographie des entreprises (date de cessation économique, date de cessation administrative, date de création, date de réactivation), compte-tenu de la période écoulée entre la date à laquelle le redressement a été réalisé (dernier trimestre 2003) et la période de collecte (fin de collecte notée au 2 avril 2002), se sont révélées très efficaces pour qualifier en entreprises cessées une partie des entreprises pour lesquelles on manquait d'information.

Ainsi on a pu requalifier 216 entreprises soit 30% des entreprises de la catégorie « questionnaires dits envoyés » en entreprises cessées.

Par ailleurs, sur les 1 671 questionnaires disponibles, 39 ont fait l'objet lors d'une étape précédente - celle du redressement de la non réponse partielle - d'une qualification en non réponse (car la qualité de remplissage des questionnaire était trop mauvaise). Il reste donc réellement 1 570 questionnaires exploitables (65 % des 2 399).

Nous obtenons donc la décomposition suivante :

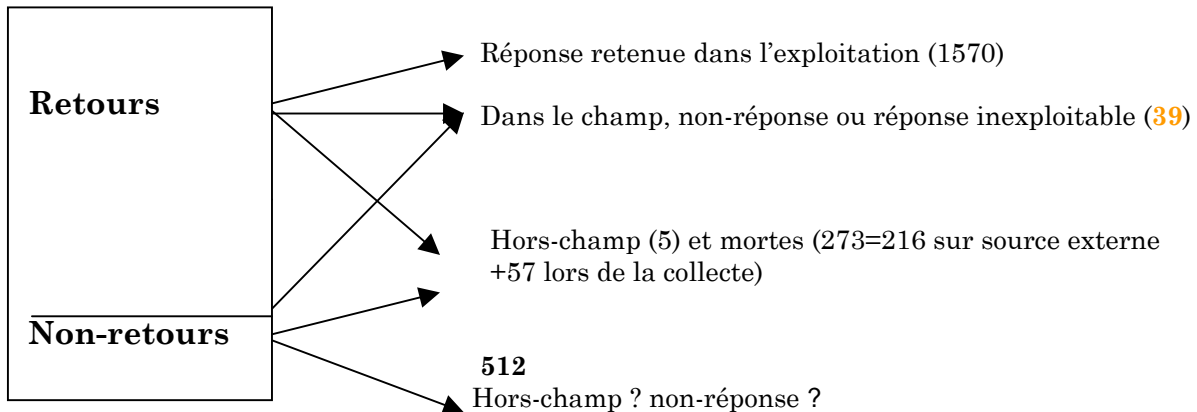


Figure 4 : Répartition des questionnaires dans les quatre catégories.

3.3. Constitution du fichier de données qui sera « calé »

Une fois cette étape de requalification des questionnaires réalisée, il est aisé de séparer les enregistrements codés en retour (ici donc 1 848 questionnaires) par rapport aux autres (551). Notons que nous avons simplifié le processus de correction de la non-réponse présenté précédemment en deux étapes en assimilant les 39 questionnaires dans la catégorie « non-réponse » aux questionnaires sans information (du fait de leur petit nombre), ce qui nous a permis de ne réaliser qu'une seule étape. Cette démarche équivaut à une petite perte d'information puisqu'on va en une seule opération augmenter le poids des 1 848 entreprises (catégories 1 et 3) pour compenser les unités de la 4^{ème} catégorie (on fait l'hypothèse qu'il n'y a pas de questionnaires inexploitable (catégorie 2 = 0)).

On crée la variable binaire (retour/non retour). C'est sur cette variable que sont fondés ensuite les traitements de la non réponse par estimation d'une régression logistique rendant le mieux compte possible des variables expliquant la probabilité de retour. Après plusieurs essais, le modèle retenu s'appuie sur le secteur d'activité et la région des entreprises, la variable taille n'introduisant pas d'amélioration du modèle.

Le poids final attribué aux entreprises en retour est le suivant :

$$\text{Poids final} = (\text{poids initial}) / \text{probabilité de retour}$$

Au final, les 1 570 entreprises dont le questionnaire est exploitable pèsent pour 82 % du total. Les mortes et les hors champ ont un poids non nul.

Tableau 1 : Résultats après redressement de la non-réponse totale

	nombre de questionnaires	En %	Nombre estimé d'entreprises	en %
	2399	100,0	285730	100,0
Hors champ	5	0,2	874	0,3
Mortes	273	11,4	49886	17,5
Statut indéterminé	551	23,0		
Répondants dans le champ	1570	65,4	234969	82,2

4. Conclusion

Quelles conclusions tirer de cette présentation, au-delà des questions méthodologiques soulevées ?

Sans doute des consignes relatives à l'organisation du traitement d'une enquête. La nécessité de garder trace du plan de sondage bien sûr, mais également de conserver le fichier contenant l'ensemble des unités sélectionnées ainsi que « l'état » de la base de sondage au moment du tirage, est ici évidente. Il s'agit donc d'insister ici sur le travail d'archivage indispensable à effectuer pour le traitement des questions méthodologiques en aval de la collecte.

Dans ce papier, nous avons également présenté une des principales difficultés méthodologiques auxquelles se heurte le statisticien pour prendre en compte de façon correcte la non-réponse, qui est la difficulté à « fabriquer » ex post les catégories qui vont être utilisées pour ce travail. L'idéal serait de prévoir dès le départ des codes de gestion adaptés à cette question.

Ajoutons que le contexte est différent pour les enquêtes auprès des ménages ; celles-ci, réalisées en grande majorité en face à face, bénéficient de « l'aide » des enquêteurs pour séparer, parmi les logements sélectionnés qui ne répondent pas, ceux qui font partie du champ de l'enquête et qui auraient dû répondre à l'enquête de ceux qui n'en font pas partie. La catégorie problématique est donc vide et le processus de correction de la non-réponse plus simple à réaliser.

Bibliographie

- [1] Berret P., Pietri-Bessy P., 2004, « Les entreprises des services innovent aussi », Insee-première n°982
- [2] Caron N., 1996, « les principales techniques de correction de la non-réponse et les modèles associés », n°9604, *document de travail de la série Méthodologie Statistique, Insee* (version actualisée disponible courant 2005).
- [3] Caron N., 2002, « Les problèmes de calage dans les enquêtes entreprises », *Présentation aux JMS de décembre 2002*.
- [4] Deville J.-C. et Särndal C., 1992, « Calibration estimators in Survey Sampling », *JASA*, vol. 87, n° 418, p. 376-382.
- [5] Rivière P., 1996, « Enquêtes annuelles d'entreprise : à la rencontre du 4^{ème} type », *courrier des statistiques* n°78, Insee.