

LES STATISTIQUES SUR LES MOUVEMENTS DE MAIN-D'ŒUVRE : PASSAGE D'UNE ENQUÊTE À UNE SOURCE ADMINISTRATIVE

Muriel BARLET¹ (*), Philippe RAYNAUD² (*) et Olivier SANZERI³ (*)

(*) Dares, sous-direction Emploi et marché du travail, département Emploi

Résumé

La déclaration sociale nominative (DSN) se substitue progressivement à la plupart des sources administratives sur l'emploi et les salaires. La Dares a été la première direction de la statistique publique concernée par ce projet, la DMMO (déclaration mensuelle des mouvements de main-d'œuvre) étant la première déclaration substituée par la DSN. En cible, la DSN se substituera également aux BRC, aux RMM et à la DADS, sources sur lesquelles reposent en grande partie les statistiques conjoncturelles et structurelles d'emploi et de salaires. Cette étude présente les difficultés rencontrées et les choix retenus par la Dares pour reconstituer des statistiques sur les mouvements de main-d'œuvre (dénombrement des contrats conclus ou rompus par motif) à partir de la DSN. Tout d'abord, les spécificités des données de paie, dont la DSN est issue, ont conduit la Dares à constituer, préalablement à la reconstitution proprement dite des mouvements de main-d'œuvre, une base des contrats historisée à partir des flux DSN. L'enjeu est notamment de réaffecter correctement les données aux périodes réelles d'activité, plutôt qu'aux périodes de paie, afin que tous les rappels de salaires et tous les « décalages » ou « retard » de paie soient comptabilisés de façon cohérente avec les flux de main-d'œuvre. Dans un second temps, les gestionnaires de la Dares pourront corriger manuellement les données incohérentes. Comme c'est déjà le cas pour les traitements actuels des MMO, une attention particulière sera portée aux déclarations des établissements touchés par une restructuration (cessation, changement de Siret...) afin de ne pas comptabiliser comme des embauches et des ruptures de contrat, les transferts de contrat d'un établissement employeur à un autre à la suite de ces événements. Enfin la Dares a prévu de mettre en place un traitement de la non réponse. Même si la DSN devrait être exhaustive (ne pas l'effectuer constituerait une fraude pour l'entreprise), il est probable que durant une phase de montée en régime du dispositif, une partie des employeurs effectueront leurs démarches avec retard, ce qui peut créer des biais de non-réponse pour les statistiques trimestrielles. Deux méthodes de redressement sont envisagées : calculer les évolutions sur les établissements présents/présents (cylindrage) ou une méthode plus usuelle de type méthode du ratio séparé.

Abstract

Text (de 5 à 10 lignes)

Mots-clés

DSN, Données administratives, flux, redressements.

¹ Muriel.barlet@travail.gouv.fr

² Philippe.raynaud@travail.gouv.fr

³ Olivier.sanzeri@travail.gouv.fr

Introduction

La déclaration sociale nominative (DSN) va progressivement se substituer aux quatre sources administratives utilisées par la statistique publique pour établir les statistiques d'emploi salarié : la DADS (Déclaration annuelle de données sociales, utilisée principalement pour les estimations d'emploi structurelles), les BRC (Bordereaux récapitulatifs de cotisations, utilisés principalement pour les estimations conjoncturelles d'emploi salarié), les RMM (relevés mensuels de mission, utilisés pour les estimations d'emplois intérimaires, lesquelles alimentent les estimations d'emploi salarié) et la DMMO/EMMO (déclaration mensuelle sur les mouvements de main-d'œuvre/enquête sur les mouvements de main-d'œuvre utilisées pour mesurer les flux de main d'œuvre). De par le calendrier de montée en charge de la DSN, la DMMO et l'EMMO sont depuis avril 2013, les premières déclarations substituées par la DSN, pour les entreprises volontaires pour entrer dans ce dispositif. La Dares, en tant qu'organisme collectant la DMMO et l'EMMO, a ainsi la charge de reconstituer les mouvements de main-d'œuvre (MMO, embauches et ruptures de contrat) à partir des données issues de la DSN. Cette étude vise à exposer les questions soulevées par cette reconstitution, et les méthodes retenues ou pressenties par la Dares pour y répondre.

Les traitements qui seront réalisés par la DARES au terme du projet actuellement en cours se répartissent en trois phases. La première est la reconstitution automatique des embauches et des ruptures de contrats à partir des données de la DSN. La deuxième consiste à apporter des corrections manuelles à ces données reconstituées automatiquement, pour traiter les anomalies dues au non-respect ou à l'incompréhension de certaines règles de la norme NEODES⁴. Ces corrections manuelles seront apportées par les gestionnaires de la DARES. La dernière phase est celle du traitement de la non-réponse, qui permettra de corriger les biais éventuels pour les statistiques de court terme dus aux déclarations retardataires ou aux possibles incidents techniques de transmission des données. Ces trois phases de traitement sont décrites successivement dans la suite de cette étude.

L'implémentation de ces traitements par la DARES est réalisée de façon progressive : la reconstitution automatique est en cours de développement ; l'outil pour les gestionnaires est en phase de conception ; le traitement de la non-réponse en est encore au stade de la réflexion. Certains des traitements décrits ci-après peuvent encore évoluer, ne serait-ce que parce que les choix définitifs ne pourront être arrêtés que lorsqu'on disposera d'un recul suffisant sur les données de la DSN. L'objectif de cette étude est donc à ce stade à la fois de synthétiser l'expérience déjà acquise par la DARES sur le dispositif DSN en vue de la production des statistiques de MMO, de présenter de façon cohérente les solutions pressenties, et aussi, en donnant une vision d'ensemble de ce projet de reconstitution des statistiques de MMO par la DSN, de permettre les échanges qui alimenteront sa propre réflexion méthodologique.

⁴ Norme d'Echange Optimisée des Déclarations Sociales, nom officiel du dispositif de déclaration DSN.

1. Reconstitution des mouvements de main-d'œuvre

1.1. Le dispositif statistique actuel sur les mouvements de main-d'œuvre

Actuellement, les statistiques sur les mouvements de main-d'œuvre (MMO) sont réalisées grâce à la fusion de deux sources :

- Une source administrative, la déclaration sur les mouvements de main d'œuvre (DMMO) ; cette déclaration mensuelle est obligatoire (art L.1221-16 du code du travail) pour les établissements de 50 salariés ou plus (depuis 1976).
- Une enquête trimestrielle par sondage auprès des établissements employeurs (enquête sur les mouvements de main d'œuvre, EMMO), qui permet d'étendre le champ des statistiques de mouvements de main d'œuvre aux établissements de 1 à 49 salariés⁵.

Le dispositif statistique sur les MMO a pour but de mesurer les flux de main-d'œuvre (en particulier leurs évolutions), par type de mouvement (embauches en CDD, en CDI, démissions, licenciements...), catégorie de salariés (sexe, âge) et d'établissements (secteur d'activité, taille d'établissement...).

Le champ est celui des établissements du secteur concurrentiel (industrie, construction et tertiaire) de 1 salarié ou plus de la France métropolitaine. Sont en particulier exclus les administrations publiques, les collectivités territoriales, les principaux établissements publics, les établissements relevant de la Défense nationale et les établissements de travail temporaire.

1.1.1. Les données collectées

Les établissements assujettis aux MMO doivent fournir mensuellement (pour les DMMO) ou trimestriellement (pour l'EMMO) deux types de renseignements : les effectifs de l'établissement et les contrats de travail conclus, échus ou rompus dans la période en question (les mouvements).

Les effectifs de l'établissement :

L'informations sur les effectifs est renseignée sous forme d'équation comptable : l'effectif de début de période + les entrées sur la période – les sorties sur la période doit être égal à l'effectif de fin de période. L'effectif de fin de période est également ventilé selon le sexe du salarié. Enfin chaque établissement doit renseigner l'effectif intérimaire présent en fin de période.

Les mouvements :

Les mouvements correspondent aux entrées (par embauche ou transfert) et aux sorties (par rupture ou fin de contrat ou transfert). Un transfert de salarié est un changement d'établissement sans rupture du contrat de travail. Théoriquement, chaque transfert donne lieu à la déclaration d'un transfert entrant par l'établissement bénéficiaire du transfert et d'un transfert sortant pour l'établissement cédant (si les deux établissements sont assujettis aux MMO). Tous les mouvements de salariés survenus entre le premier jour de la période (mois ou trimestre) à 0h et le dernier jour de la période à minuit doivent être déclarés par l'employeur.

Pour chaque mouvement déclaré, l'établissement renseigne les caractéristiques du salarié concerné (nationalité⁶, date de naissance, sexe, PCS) et du contrat : date d'entrée et nature du contrat (CDD, CDI, transfert entrant) quel que soit le mouvement, et s'il s'agit d'une sortie, date et motif de sortie (licenciement, démission, fin de CDD ...).

⁵ L'EMMO, créée en 1989, portait à l'origine sur les établissements de 10 à 49 salariés ; elle a été élargie à ceux de 1 à 9 salariés en 2007.

⁶ Français, ressortissant de l'UE, autre nationalité.

1.1.2. Saisie, redressement et apurement des données collectées

Les réponses collectées sont transmises au prestataire de saisie des MMO. Schématiquement, les réponses par le biais des questionnaires et CERFA envoyés aux établissements sont saisies par lecture optique, alors que les réponses transmises au prestataire sur un document non standard sont saisies manuellement. Chaque année 6,3 millions de lignes sont saisies par le prestataire.

Les réponses présentant après saisie une incohérence ou des valeurs manquantes sont triées selon la gravité de cette incohérence. Les établissements concernés sont alors contactés par téléphone par une équipe de téléopérateurs pour corriger les déclarations. Chaque année 220 000 incohérences sont ainsi corrigées dans les équations comptables et 590 000 le sont sur les déclarations de mouvements.

Pour rendre ces données exploitables statistiquement, la DARES redresse les anomalies résiduelles (établissements non répondants, réponses présentent des incohérences ou valeurs manquantes) via une application statistique dédiée (redressement de la non-réponse, partielle et totale).

1.2. La déclaration sociale nominative

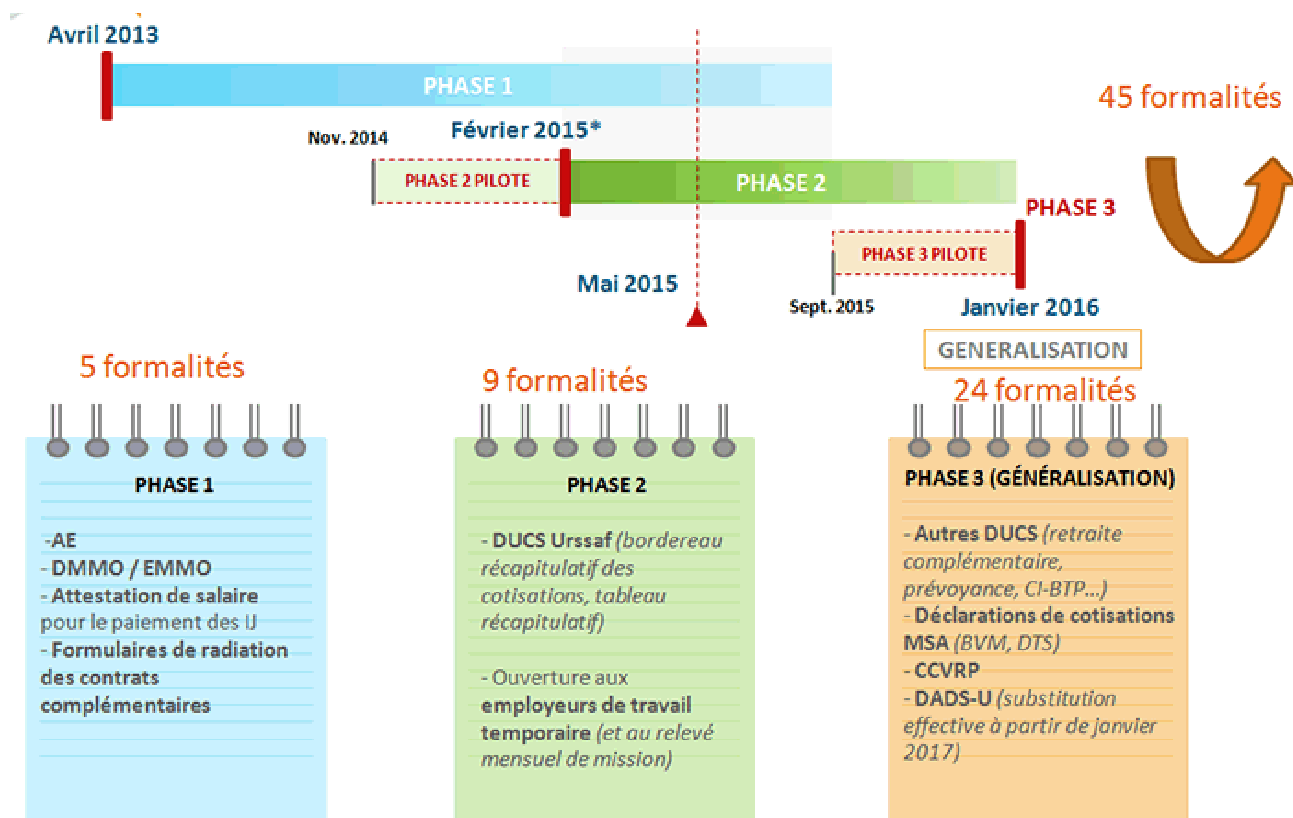
1.2.1. Généralités

La particularité de la DSN est son intégration aux systèmes d'information de gestion des ressources humaines (en premier lieu les logiciels de paie) des employeurs : les données de gestion saisies par les entreprises alimentent directement et de façon automatique le système DSN, qui va ensuite alimenter, également de façon automatique, les organismes utilisateurs de ces données. Pour coordonner ce système complexe, les données sont structurées de façon précise et cohérente par une norme technique, qui en définit les concepts et les formats. Cette norme, décrite dans le « cahier technique » permet d'assurer de façon cohérente la reconstitution des différentes formalités administratives auxquelles la DSN se substitue, en minimisant le nombre de variables collectées. Une conséquence de cette intégration de la DSN dans les SI de gestion est que les éditeurs de logiciels de paie sont des acteurs importants de la DSN, à l'interface des employeurs (« déclarants ») et des organismes utilisateurs des données de la DSN. La DSN est ainsi souvent présentée comme la « dernière opération de la paie ».

La DSN va progressivement remplacer l'ensemble des déclarations sociales. Le périmètre déclaratif de la DSN s'élargit progressivement en incluant de nouvelles déclarations et de nouvelles entreprises cibles. Son déploiement a été organisé en 3 phases :

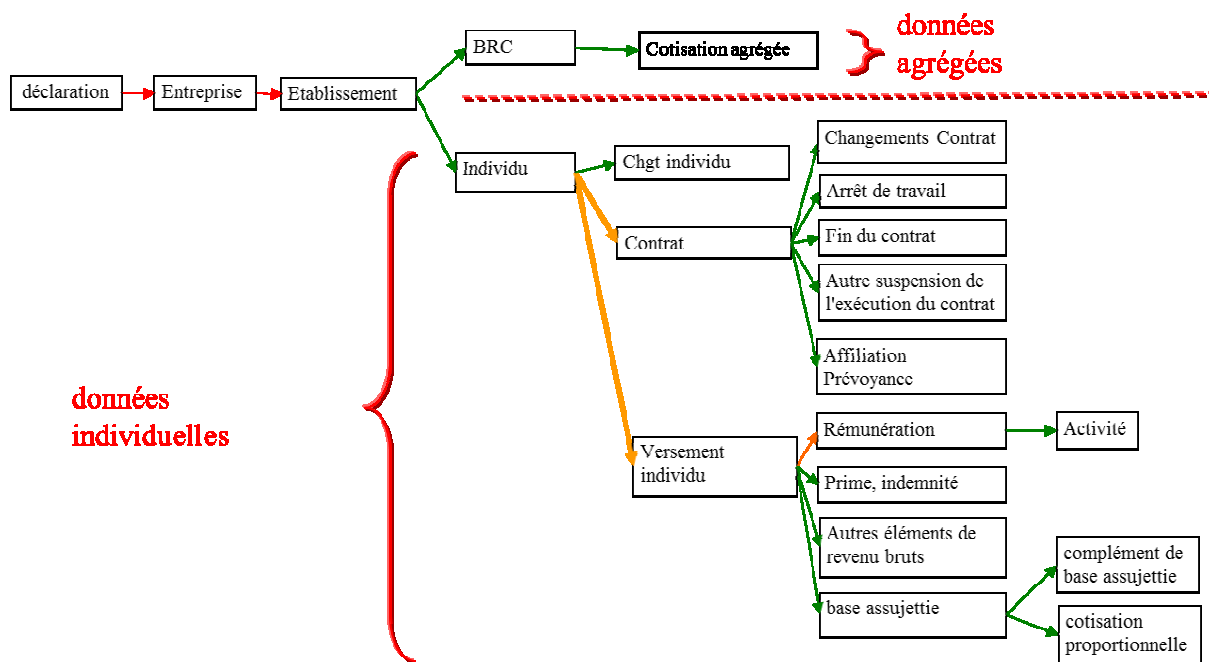
Le déploiement du système DSN :

CALENDRIER DSN



Les données présentées plus bas sont celles de la phase 2.

La DSN est une déclaration complètement dématérialisée. Tous les mois, les établissements doivent transmettre un fichier hiérarchisé contenant des informations sur l'établissement et l'entreprise, les cotisations dues, les salariés, les contrats de travail de chaque salarié et les rémunérations (voir schéma ci-dessous).



Les valeurs contenues dans les rubriques de cette DSN « mensuelle »⁷ sont celles en vigueur le dernier jour du mois. Les modifications de certaines variables en cours de mois sont renseignées par un système de « blocs changements », qui joue en particulier un rôle essentiel pour les variables identifiantes des contrats de travail en permettant le chaînage des observations, et donc le suivi des contrats de travail au cours du temps. Les « blocs changements » renseignés par les employeurs doivent comporter les valeurs précédentes des variables qui ont été modifiées en cours de mois, et les dates de changement. Ainsi, si un salarié voit son CDD transformé en CDI le 10 du mois, la rubrique « nature de contrat » prendra comme nouvelle valeur « CDI » et la rubrique « ancienne nature de contrat » prendra la valeur « CDD », la date de modification associée étant le 10 du mois.

Les données de la DSN mensuelle nécessaires à la reconstitution des MMO (cf. 1.1 ci-dessus) sont :

- Pour l'ensemble des salariés : les dates de début et de fin de contrat (qui permettent de déterminer l'effectif présent en début et en fin de mois), ainsi que le sexe.
- Pour ceux concernés par un mouvement : la nature du contrat de travail, la PCS, la nationalité, l'ancien Siret d'affectation (qui permet de repérer les transferts de contrat de travail).
- Pour les salariés des entreprises de travail temporaire en contrat de mission (c'est-à-dire les intérimaires) : le lieu de travail, qui renseigne sur l'établissement dans lequel l'intérimaire effectue sa mission. Cette information permet de reconstituer les effectifs intérimaires présents en fin de mois dans les établissements « utilisateurs », à partir des DSN déclarées par les établissements de travail temporaire⁸ (voir infra).

L'existence d'un cadre formalisé (norme technique) pour la structure des données renseignées en DSN n'implique pas une homogénéité des réponses des entreprises. Leurs comportements de déclaration va en effet refléter en partie leurs pratiques de paie antérieures, qui peuvent être très disparates d'une entreprise à une autre. Deux cas importants ayant un impact sur la reconstitution des MMO sont de ce point de vue les décalages de paie et les rappels de salaire. Par ailleurs, un des principes de la DSN, combinée au projet « dites-le nous une fois », est qu'une information qui peut être reconstituée n'est pas demandée directement dans la déclaration. Il en va ainsi des embauches,

⁷ Il existe également une DSN dite « événementielle » qui permet de décrire, au fil de l'eau des événements affectant les données individuelles, une rupture de contrat de travail par exemple. Ces événements sont ensuite réintégrés dans la DSN mensuelle. La Dares utilise uniquement cette dernière.

⁸ Rappelons que la DSN se substitue également à terme aux Relevés mensuels de mission, formalité administrative obligatoire pour les ETT.

des transferts entrants et des effectifs salariés et intérimaires qui peuvent théoriquement être recalculés à partir des données de la DSN, et ne sont donc pas directement déclarés. Le système d'information conçu par la Dares pour reconstituer les mouvements de main-d'œuvre vise à traiter automatiquement ces deux difficultés : le traitement des spécificités de la paie et la reconstitution des variables non directement déclarées.

1.2.2. Spécificité des données de paie

La DSN reflète la paie du mois de référence auquel elle se rapporte, mais pas nécessairement de façon exacte l'activité de ce mois. Il est ainsi fréquent que la paie soit calculée quelques jours avant la fin du mois, de sorte que les événements de fin de mois n'y apparaissent pas : un salarié embauché en toute fin de mois risque de n'apparaître que dans la DSN du mois suivant, alors que pour l'analyse des mouvements de main-d'œuvre cette embauche devrait bien être comptabilisée au mois de sa réalisation. Le traitement des données DSN doit donc permettre d'affecter les événements qu'elles permettent de reconstituer (embauches, fins de contrat, changement de nature de contrat...) non pas à leur mois de déclaration mais à leur mois de réalisation. Il en résulte que d'une part les mouvements de main-d'œuvre d'un mois M seront calculés non seulement à partir de la DSN du mois M mais également à partir des DSN du ou des mois suivants, l'ampleur du décalage maximum entre la réalisation d'un événement et son apparition en DSN n'étant a priori pas connue.

Par ailleurs, la présence de rappels de salaire (c'est-à-dire de versements relatifs à un mois donné, mais effectués ultérieurement) sera probablement relativement fréquente. Le risque est que ces rappels, lorsqu'ils portent sur un contrat échu, soient interprétés de façon erronée, comme de nouveaux mouvements de main-d'œuvre. Pour éviter de telles confusions, il est nécessaire, lorsqu'un contrat apparaît dans la DSN du mois M (au sens où son identifiant n'était pas présent dans la DSN du mois M-1), de vérifier que ce contrat n'avait pas déjà été déclaré dans une DSN antérieure. S'il s'agit bien d'un nouveau contrat en DSN, on traitera cette observation comme une embauche. Dans le cas contraire, on considèrera qu'il s'agit d'un rappel de paie et que les MMO associés à ce contrat ont déjà été pris en compte.

Ainsi, du fait des spécificités des données de paie, la reconstitution des MMO pour un mois M à partir des données DSN nécessitent de mobiliser les informations contenues dans les DSN suivantes (pour repérer les événements signalés tardivement) mais également dans les DSN précédentes (pour repérer les contrats présents uniquement au titre de rappel de salaire).

1.2.3. Reconstitution des éléments non déclarés directement en DSN

Les embauches et les transferts sortants ne sont pas déclarés en tant que tels en DSN. Il faut les reconstituer à partir des données déclarées.

S'agissant des embauches, aucune rubrique ne permet de savoir qu'un contrat de travail est déclaré pour la première fois. La DSN mentionne en revanche bien la date de début de contrat de travail. Mais pour les raisons développées ci-dessus (retard de déclarations, rappels de salaires) le repérage des embauches nécessite à la fois de vérifier systématiquement que le contrat est déclaré pour la première fois en DSN, et d'exploiter l'information sur la date de début.

Les transferts sortants sont également absents des déclarations DSN. Seuls les transferts entrants sont déclarés et l'établissement bénéficiaire a la charge de déclarer l'établissement d'origine du salarié transféré. Ainsi il est nécessaire d'utiliser les informations contenues dans la déclaration de l'établissement bénéficiaire pour reconstituer l'ensemble des mouvements de l'établissement cédant.

Enfin, l'effectif intérimaire en fin de mois ne peut pas être une information contenue dans la DSN. En effet, par définition les intérimaires ne sont pas payés par l'établissement utilisateur mais par l'entreprise de travail temporaire. Aussi, aucune information sur cette population n'est présente dans les bases de données des logiciels de paie qui émettent la DSN. Cette information ne peut être reconstituée qu'à partir des déclarations des établissements de travail temporaire. Cela nécessite toutefois qu'ils déclarent, avec une qualité suffisante, le lieu de travail de leurs salariés en mission (l'identité de l'établissement utilisateur). Dans les déclarations substituées par la DSN (le relevé

mensuel de mission, notamment), cette variable n'est renseignée que pour la moitié des missions environ. La complétude et la qualité de cette information devrait donc s'améliorer avec la mise en place de la DSN.

Afin de réaliser de façon cohérente et efficace ces reconstitutions, la Dares a choisi de réorganiser les données DSN dans une base historisée des contrats de travail, puis de calculer les mouvements de main-d'œuvre à partir de cette base.

1.3. Base des contrats de travail et calcul des MMO

1.3.1. La base des contrats de travail

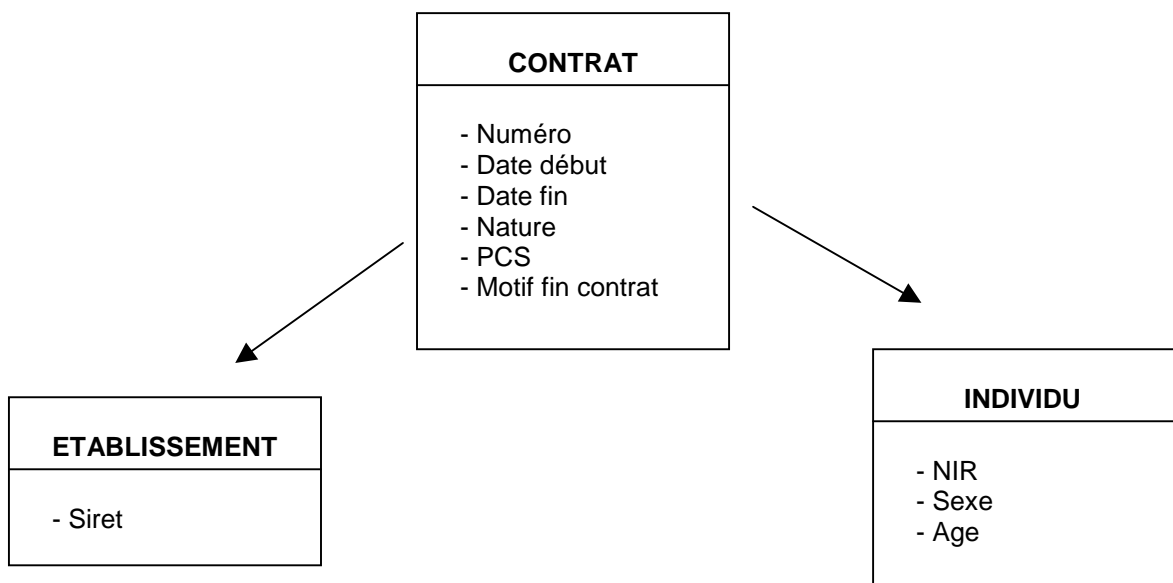
La base des contrats de travail est organisée autour de trois concepts : l'établissement, l'individu et bien sûr le contrat de travail. Chacun de ces concepts a des attributs qui peuvent ou non être modifiés par les informations contenues dans les DSN successives. Un contrat de travail est rattaché à un individu et un établissement. Un contrat de travail donné est toujours rattaché au même individu mais il peut changer d'établissement d'affectation (lors des transferts).

Schématiquement, la base des contrats de travail est initialisée à partir des premières données DSN. Elle est ensuite alimentée chaque mois avec les dernières données DSN (reçues à la Dares le 18 du mois). Les contrats contenus dans ces nouvelles DSN sont systématiquement recherchés parmi ceux déjà déclarés en DSN, par appariements avec la base des contrats historisée, grâce à l'identifiant du contrat de travail DSN (constitué en DSN à partir de l'établissement d'affectation, du NIR, de la date de début de contrat et du numéro de contrat). Si l'appariement est réussi et le contrat par conséquent déjà présent dans la base historisée, les informations contenues dans cette dernière sont mises à jour à partir des changements déclarés en DSN (modification de la nature de contrat, de la PCS...). Si l'appariement échoue un nouveau contrat de travail est ajouté dans la base historisée.

Cette méthode permet de distinguer les rappels de salaires des embauches : la comparaison systématique des contrats contenus dans les nouvelles DSN et de ceux contenus dans la base historisée permet de repérer les contrats qui ont déjà été signalés (cas des rappels). Elle permet également de repérer les transferts sortants, à travers l'historisation des établissements d'affectation au niveau de chaque contrat. Enfin, les mouvements de main-d'œuvre sont calculés à partir de la base des contrats et des dates d'effet déclarées (date de début et de fin de contrat) en s'affranchissant totalement du mois de référence de la DSN par laquelle l'information est arrivée. De la sorte, tous les événements sont affectés à leur mois de réalisation et non pas à leur mois de déclaration.

En pratique, le processus d'intégration d'une nouvelle DSN à la base des contrats de travail est un peu plus complexe. L'établissement d'affectation, les individus puis les contrats sont successivement recherchés (par appariement sur leurs identifiants) dans la base des contrats. À chaque étape, si l'appariement échoue une nouvelle « ligne » est insérée dans des tables établissements, individus ou contrats. Ce processus est détaillé ci-dessous.

Base des contrats :



L'identifiant de l'établissement d'affectation est le SIRET. L'appariement des nouvelles DSN avec la base des contrats sur cet identifiant permet d'actualiser les informations sur les établissements contenues dans la DSN (code APET, code postal, localité, effectif de fin de mois). Pour chaque établissement, on calcule également le mois de première et de dernière DSN. Le mois de première DSN sert essentiellement à savoir à partir de quel mois il faut reconstituer la DMMO ou l'EMMO pour cet établissement. Le mois de dernière DSN permet de vérifier que les déclarations sont intégrées dans l'ordre ou de repérer les déclarations manquantes. Tant qu'un établissement n'est pas cessé, une DSN est attendue si l'établissement a précédemment effectué une DSN. En effet, mêmes les établissements temporairement sans salarié doivent émettre une DSN. Il existe un type de DSN pour ce cas de figure particulier : les « DSN néants ».

Au niveau de l'individu, on réalise l'appariement sur la base du NIR certifié si celui-ci est présent, c'est-à-dire dans les cas où la CNAV, à partir des informations transmises dans la DSN, sera parvenue à certifier l'identité de l'individu. En l'absence de NIR certifié, l'appariement est testé sur la base du NTT (numéro technique temporaire). Le NTT est un numéro fourni par l'entreprise aux salariés qui ne sont pas en mesure de communiquer un NIR. Ce numéro contient le SIREN et le matricule du salarié, il est donc stable pour un même individu dans une même entreprise. Le mois où le déclarant connaît le NIR du salarié, il renseigne dans la DSN le NIR et le NTT ce qui permet de chaîner les déclarations successives. Enfin en l'absence de NIR et de NTT un appariement sur la base du SIRET de l'établissement d'affectation, du nom de famille, du prénom et de la date de naissance est testé. Ce cas de figure peut se produire dans le cas où un NIR est renseigné (ce qui lève l'obligation de renseigner un NTT) dans la déclaration mais où la CNAV ne parvient pas à certifier l'identité. Par ailleurs dès qu'un bloc changement est renseigné pour une des variables utilisées dans l'appariement c'est la valeur contenue dans le bloc changement qui est utilisée pour l'appariement. En effet, cette valeur correspond à la valeur connue dans la DSN précédente, c'est donc sur celle-là que l'appariement doit se faire. Enfin, les attributs de l'individu reportés dans la base des contrats sont : le sexe, la nationalité, le nom de famille et le nom d'usage, les prénoms, la date de naissance.

Pour le contrat de travail, enfin, l'appariement est effectué sur la base des identifiants de l'établissement d'affectation et de l'individu, ainsi que de la date de début de contrat et du numéro de contrat (utile dans le cas où un salarié a plusieurs contrats simultanés avec le même employeur). Si un bloc changement est renseigné pour l'une de ces variables, par exemple, un ancien numéro de contrat de travail, c'est la valeur contenue dans ce bloc changement qui est retenue pour l'appariement (cette valeur permet de chaîner les informations du mois courant avec celles des mois précédents, en vue de l'intégration dans la base des contrats). Les attributs du contrat qui sont stockés sont : le libellé d'emploi, la date de fin prévisionnelle, la nature du contrat, le code de régime de base maladie, la PCS. Pour tous ces attributs un historique est conservé ce qui permet de connaître ces caractéristiques à n'importe quel moment de la vie du contrat. Pour la reconstitution des mouvements de main-d'œuvre il est, en effet, a minima nécessaire de connaître les caractéristiques du contrat au moment de l'embauche, plus spécifiquement la nature du contrat (CDD ou CDI) et la PCS du salarié qui l'occupe. Les attributs du contrat ne sont actualisés que lorsque la modification est déclarée via un bloc changement, qui est la manière correcte de déclarer les évolutions en DSN pour que celles-ci puissent être datées précisément. Ainsi, les changements non déclarés ne sont pas enregistrés. Une exception est faite pour la nature de contrat (variable centrale des mouvements de main-d'œuvre) pour laquelle une comparaison systématique est effectuée entre l'information contenue dans la base des contrats et l'information contenue dans les nouvelles DSN. En plus des caractéristiques, on conserve pour le contrat le mois de la dernière déclaration dans laquelle le contrat est apparu. Cela permet de vérifier qu'un contrat pour lequel aucune date de fin de contrat n'a été déclarée continue bien d'être renseigné dans les DSN successives.

Une attention particulière est portée au changement d'établissement d'affectation. En effet, lorsqu'un ancien établissement d'affectation est présent dans une DSN, cela signifie que le contrat a été transféré d'un établissement à un autre, sans rupture du contrat. Il est donc nécessaire de clore le lien entre le contrat et l'ancien établissement d'affectation (alors même que le transfert n'est déclaré que par le nouvel établissement) et de lier le contrat au nouvel établissement.

Quelques cas qui sont considérés comme des anomalies ne sont pas traités automatiquement. Une fois la première DSN pour un établissement donné intégrée dans la base des contrats, si une déclaration correspond à une date antérieure est reçue, alors elle n'est pas intégrée. De même, le système n'intègre la DSN d'un établissement que si la DSN immédiatement antérieure a été reçue, sinon la DSN est mise en attente. Enfin, lorsque plusieurs anciens établissements d'affectation sont renseignés, ce qui n'est théoriquement possible qu'en cas de correction sur une déclaration de changement d'établissement (correction de date ou d'ancien Siret) ou « d'aller-retour » d'un salarié entre deux établissements au cours du même mois, le contrat est retourné en anomalie.

1.3.2. La reconstitution des MMO

Les mouvements de main-d'œuvre sont ensuite calculés à partir de la base des contrats ainsi constituée. Ils sont définis de la manière suivante :

- Une embauche est générée pour chaque contrat dont la date de début appartient au mois d'intérêt et à la condition que la date de première DSN pour l'établissement d'affectation du contrat soit postérieure ou égale au mois d'intérêt (on ne génère pas de mouvements antérieurs au mois d'entrée d'un établissement en DSN). L'embauche est rattachée à l'établissement d'affectation du contrat à sa date de début. La nature de contrat et la PCS du salarié sont celles en vigueur à la date de début de contrat. Concrètement, si un contrat conclu en CDD est transformé en CDI, l'embauche reconstituée doit être en CDD quelle que soit la date de changement de la nature de contrat.
- Une fin de contrat est générée pour chaque contrat dont la date de fin est présente et dans le mois d'intérêt. La fin de contrat est caractérisée par le motif de fin de contrat, la nature du contrat et la PCS du salarié à l'échéance du contrat. La fin de contrat est rattachée à l'établissement d'affectation du contrat à son échéance.
- À chaque changement d'établissement d'affectation d'un contrat, un transfert entrant et un transfert sortant sont générés. Pour les deux types de transfert, on conserve l'établissement « partenaire ». Cette information paraît pertinente à stocker pour étudier ex-post à partir des

données générées les restructurations d'entreprises à partir des mouvements groupés de salariés.

Les données de la base des contrats permettent également de calculer l'effectif en fin de mois, en comptabilisant pour un établissement l'ensemble des contrats en cours le dernier jour du mois. Cet effectif est décliné par sexe. Enfin, à partir des déclarations dans les DSN des établissements de travail temporaire du lieu de travail des salariés en contrat de mission (autrement dit les intérimaires), le nombre d'intérimaires présents le dernier jour du mois par établissement utilisateur est calculé. Cette statistique ne sera pertinente que lorsque tous les établissements de travail temporaire effectueront une DSN.

Lors de la constitution de la base des contrats un certain nombre d'anomalies seront détectées. Une partie d'entre elles sera corrigée manuellement par des gestionnaires (cf. partie suivante), d'autres seront corrigées par des méthodes d'imputation statistique (cf. partie 3)

2. Correction des anomalies

2.1. Rappel de l'existant

Dans le cadre de la collecte actuelle de la DMMO/EMMO, deux grands types d'anomalies sont traités : les incohérences et les anomalies « référentiel ».

Les incohérences regroupent : les problèmes de bouclage de l'équation comptable (entrées – sorties différents de l'écart entre l'effectif de début de mois et l'effectif de fin de mois), les écarts entre les mouvements et l'équation, le non-respect des nomenclatures (par exemple pour la PCS). Ces anomalies sont corrigées par le prestataire de saisie des mouvements de main-d'œuvre suivant des priorités fixées par la DARES. Les incohérences qui n'ont pas pu être corrigées lors de l'exploitation statistique des données sont redressées par des méthodes d'apurement/imputation.

Les anomalies référentiel regroupent : les informations relatives à la démographie des établissements reportées sur les formulaires par les déclarants (cessation, changements de SIRET) et les « risques » c'est-à-dire les évolutions d'effectifs ou de mouvements laissant présager une restructuration des établissements (passage à un effectif nul, forte augmentation de certain type de mouvement, évolution exceptionnelle des effectifs...). Il est important pour la qualité des données sur les mouvements de main-d'œuvre de traiter correctement ces anomalies référentiel. D'une part, le suivi des changements de SIRET permet de prévenir l'attrition de l'échantillon en cours d'année.. D'autre part, il est important de distinguer les mouvements de main-d'œuvre « réels », traduisant des événements d'embauche ou de ruptures de contrats de travail, de mouvements de main-d'œuvre apparents résultant uniquement de restructurations d'entreprises : par exemple, lorsqu'un établissement change de SIRET, tous les contrats de ses salariés sont automatiquement rattachés au nouveau SIRET sans pour autant qu'aient eu lieu autant d'embauches pour le nouveau Siret et de fins de contrat de travail pour l'ancien. La gestion des anomalies référentiel est effectuée par une équipe de gestionnaire de la DARES.

Grâce à la DSN, en régime permanent,, certaines de ces anomalies devraient disparaître. En effet, comme la déclaration est complètement dématérialisée, il sera possible de contrôler automatiquement la validité des réponses (par exemple en termes de cohérence avec les nomenclatures ou d'obligation de renseigner certains champs). Le dépôt de la déclaration peut alors être bloqué, ce qui empêche l'employeur de se libérer de ses obligations déclaratives. Les données DSN ne devraient donc pas contenir de non réponse partielle si ce principe est appliqué strictement pour garantir la conformité des déclarations à la norme. En revanche, même s'il n'y a pas d'enjeu de collecte pour la DARES⁹, le suivi de la démographie des établissements reste un enjeu pour la qualité des statistiques produites. Il est en effet peu probable que les déclarants renseignent correctement, dès le début de la DSN, les transferts de contrat entre établissements lors des restructurations (cela devrait être théoriquement le cas après la phase d'appropriation de la DSN par les entreprises).

⁹ L'Acoss est en charge d'effectuer les relances et d'appliquer les sanctions prévues en cas d'absence de DSN.

2.2. Les anomalies repérées dans les données DSN

La DARES étant destinataire des données issues de la DSN depuis l'origine du projet (pour les entreprises pilote à partir d'avril 2013) elle a été en mesure d'analyser les premières déclarations.

Les expertises ont porté essentiellement sur les contrats qui « disparaissent » entre deux DSN mensuelles successives sans que la DSN ne contienne une date et un motif de fin de contrat de travail. La DARES avait anticipé que la déclaration des transferts de contrat lors des mutations ou des changements de SIRET ne serait pas effectuée immédiatement et correctement par les déclarants, avec pour conséquence la « disparition » des contrats concernés entre deux DSN successives, toutes les informations nécessaires pour suivre le passage des contrats d'un établissement à un autre n'étant pas renseignées. La norme NEODES impose sur ce point une nouvelle pratique de déclaration : seul l'établissement bénéficiaire du transfert doit le déclarer, ce qui nécessite qu'il connaisse le Siret de l'établissement cédant, et renseigne ce Siret dans la rubrique « ancien établissement d'affectation ». Ainsi, la « perte » des contrats de travail de l'établissement cédant doit se déduire de la déclaration de l'établissement bénéficiaire. Ce mode déclaratif n'existait dans aucune des déclarations substituées, il est donc logique qu'un temps d'adaptation soit nécessaire pour que l'ensemble des établissements employeurs l'applique correctement.

Les analyses effectuées sur les DSN de l'année 2014 montrent en effet que les disparitions de contrats entre deux DSN successives sont nombreuses. Néanmoins l'absence de déclaration de l'ancien établissement d'affectation est loin d'être le motif principal de ces « disparitions ». Rappelons qu'en norme DSN, les identifiants du contrat de travail sont : le Siret de l'établissement d'affectation, l'identité du salarié, la date de début de contrat et le numéro de contrat. Si l'un de ces identifiant est modifié entre la DSN du mois M et celle du mois M+1, sans que cette modification ne soit signalée dans un bloc changement comme cela devrait être le cas, les contrats des deux déclarations ne peuvent pas être chaînés. Ainsi, en appliquant les règles de constitution de la base de contrats décrites dans le paragraphe précédent, le contrat contenu dans la DSN du mois M « disparaît » dans la déclaration du mois M+1 sans date de fin de contrat de travail. En 2014¹⁰, 155 000 contrats ont disparu entre deux DSN successives sur une moyenne mensuelle de 180 000¹¹ contrats présents en fin de mois (tableau 1).

¹⁰ D'après les déclarations reçues par la DARES au 16 février 2015.

Tableau 1 : nombre de disparitions de contrats entre deux DSN mensuelles successives et répartition par motif

Mois	Nombre de disparitions de contrat	Modification non déclarée de l'établissement d'affectation	Modification non déclarée de la date de début de contrat	Modification non déclarée du numéro de contrat	Autre	Effectifs présents en fin de mois
Janvier	209	25%	6%	0%	69%	52 139
Février	50	34%	22%	0%	44%	55 492
Mars	306	84%	4%	1%	12%	55 730
Avril	163	58%	9%	5%	28%	58 968
Mai	222	56%	7%	1%	36%	62 024
Juin	484	9%	11%	2%	77%	71 982
Juillet	298	33%	9%	5%	53%	79 062
Août	2 719	12%	3%	75%	9%	86 904
Septembre	3 090	66%	24%	2%	8%	136 233
Octobre	7 260	18%	43%	1%	39%	262 126
Novembre	22 467	7%	18%	13%	62%	380 004
Décembre	117 746	3%	2%	75%	20%	390 047

Champ : Déclarations relatives à 2014

Source : DSN. Calcul : Dares

Lecture : 117 746 contrats ont été déclarés en décembre 2014 et non déclarés en janvier 2015 alors qu'aucune date de fin n'était renseignée pour ces contrats. Pour 3 % de ces contrats, un contrat similaire à l'exception de l'établissement d'affectation a été déclaré en janvier 2015.

Pour la plupart des contrats qui disparaissent, on retrouve dans la déclaration précédente un contrat similaire à un identifiant près (Siret d'affectation, numéro de contrat, date de début de contrat, ou identité du salarié). Suivant les mois, une majorité de contrat est retrouvée à l'exception du Siret de l'établissement d'affectation, à l'exception de la date de début de contrat ou à l'exception du numéro de contrat. En revanche, très peu d'anomalies sont liées à la déclaration de l'identité du salarié. Le motif « autre » du tableau 1 regroupant essentiellement l'absence totale de déclaration de l'établissement ou d'autres cas non analysés encore à ce stade.

2.3. Méthode de correction des anomalies envisagée

Les anomalies décrites plus haut sont généralement concentrées sur quelques établissements, qui seront traités par les gestionnaires de la DARES, dans un premier temps essentiellement pour deux types d'anomalies : le transfert des contrats de travail sans déclaration du changement d'établissement d'affectation et les changements de numéro de contrat sans déclaration de la correction. Les anomalies liées à l'absence de signalement d'une correction de la date de début de contrat seront traitées dans un deuxième temps. Elles sont en effet plus complexes, notamment dans le cas de successions de contrats courts.

Concernant les transferts de contrat de travail, la correction sera fondée sur l'information présente ou ajoutée dans le référentiel des établissements de la base des contrats. Ce référentiel propre à la base des contrats sera dans un premier temps alimenté par le référentiel commun des déclarants (RCD) de l'Acoss qui est le référentiel associé à la DSN. La qualité de cette base ayant une incidence sur l'activité de recouvrement de cotisations sociales, on peut penser que le RCD fera l'objet d'un suivi attentif par l'Acoss. Cette base pourra en outre être complétée par les gestionnaires de la DARES. Lorsqu'un changement de Siret sera renseigné dans ce référentiel, les contrats identiques (c'est-à-dire rattachés au même individu, avec la même date de début de contrat et avec le même numéro de début de contrat) des deux Siret seront « fusionnés ». Ainsi, en renseignant un changement de Siret, un gestionnaire traitera toutes les anomalies liées au non report de ce changement au niveau de chaque contrat dans les DSN mensuelles.

Le référentiel permettra également d'autres corrections automatisées. Ainsi, comme aujourd'hui en DMMO, lorsqu'un établissement est « cessé économiquement » selon le référentiel, tous les contrats qui lui sont rattachés sont clôturés. Enfin, pour chaque établissement non cessé économiquement, ayant au moins un contrat en cours et ayant déjà effectué au moins une DSN, le référentiel d'établissement associé à la base des contrats permettra de connaître les mois pour lesquels aucune DSN n'est parvenue à la DARES. Cette information sera utilisée pour les corrections de non réponse décrites dans la partie suivante.

Les gestionnaires pourront également fusionner les contrats qui ont les mêmes identifiants à l'exception du numéro de contrat, c'est-à-dire qui dépendent du même établissement, du même individu et qui ont la même date de début. Cette correction se fera simultanément pour tous les contrats d'un établissement pour lesquels l'erreur est observée. Ainsi, comme pour les transferts non déclarés, les anomalies sont regroupées et traitées par établissement, et non au niveau de chaque donnée individuelle de contrat.

Les anomalies à corriger seront probablement trop nombreuses en regard du nombre de gestionnaires de la DARES (3,8 ETP). Afin d'effectuer les corrections les plus pertinentes possibles, ces anomalies seront priorisées. A priori, les établissements en ayant le plus grand nombre seront traités en premier. Cette priorisation sera paramétrable.

Ces premières corrections manuelles seront suivies de traitements statistiques des cas d'absence totale de DSN pour un établissement actif auxquels des contrats en cours sont rattachés. Pour les établissements sans salarié, même si théoriquement une DSN néant est attendue, aucun redressement ne sera effectué. En effet, pour ces cas là, la meilleure correction serait probablement d'imputer une déclaration néant (sans salarié ni mouvement). Or, comme aucune des statistiques produites par la DARES sur les mouvements de main-d'œuvre ne dépend du nombre d'établissement employeur, ne pas traiter ces cas est équivalent à imputer une déclaration néant.

3. Redressement de la non réponse

3.1. Méthode actuelle en MMO

La non-réponse est dite totale lorsque les établissements n'ont pas renvoyé de questionnaire EMMO ou de déclaration sur les mouvements de main d'œuvre. Un établissement peut être non répondant en raison d'un retard dans l'envoi de sa réponse. Un établissement n'ayant pas répondu au moment de l'exploitation pour la période t est redressé de la non-réponse. Si cet établissement répond ultérieurement, sa réponse sera prise en compte lors de la révision des données. En effet, la DARES met à jour les données MMO tout au long de l'année en cours. Par exemple, lors de la publication du 3^{ème} trimestre 2014, les données des 1^{er} et 2^{ème} trimestre sont mises à jour (prise en compte des retardataires). Le phénomène de non-réponse est aussi bien présent dans l'EMMO que dans la DMMO.

3.1.1. Imputation des DMMO manquantes en amont du redressement de la non-réponse totale

Les périodicités de réponse diffèrent pour la DMMO et l'EMMO (mensuelle et trimestrielle). Cependant, la DARES publie des indicateurs statistiques sur la source fusionnée des MMO. Les données issues de la DMMO sont donc trimestrialisées. Un établissement issu de l'EMMO est dit non-répondant s'il n'a pas répondu au trimestre T. En revanche, un établissement assujéti à la DMMO est considéré comme non-répondant (total) s'il n'a répondu pour aucun des 3 mois du trimestre T. Dans le cas où seules une ou deux réponses mensuelles sont disponibles, les mois manquants sont imputés pour conserver ces réponses partielles.

La méthode utilisée pour cela est celle d'un *hot-deck* métrique : pour chaque établissement répondant avec des mois manquants (un receveur) on cherche l'établissement répondant sur tout le trimestre (le donneur) et appartenant à la même région, au même secteur (en A38 de la NAF REV2) et la même tranche de taille (1 à 9 salariés, 10 à 49, 50 à 99, 100 à 249 et 250 salariés ou plus), lui « ressemblant » le plus en termes de mouvements de main d'œuvre pour les mois communs de déclaration. Les mouvements des mois manquants du receveur sont alors imputés par ceux du donneur.

Pour trouver le donneur le plus « proche » d'un receveur, on définit de la façon suivante la distance entre l'établissement receveur et les établissements donneurs de la même région, du même secteur et de la même tranche de taille que le receveur :

$$DIST(\text{receveur}, \text{donneur}) = \sqrt{\sum_{\substack{\text{mois } i \text{ disponibles} \\ \text{du receveur}}} \sum_{\substack{\text{types de} \\ \text{mouvement } j}} (r_{ij} - d_{ij})^2},$$

avec

r_{ij} le nombre de mouvements du receveur le mois i et de type j (embauche en CDD, démission...)

d_{ij} le nombre de mouvements du donneur le mois i et de type j .

Les mouvements manquants d'un receveur sont imputés par ceux de l'établissement avec lequel la distance DIST est la plus faible (le donneur). Si un receveur ne trouve pas donneur dans le même croisement de région, de secteur et de taille, on en recherche un dans le même secteur et la même tranche de taille. Faute de cela, les réponses du receveur sont supprimées et il sera redressé de la non-réponse totale (voir ci-dessous « Correction de la non réponse totale par repondération »).

Une fois les mois manquants imputés, des tests sont réalisés afin de vérifier la cohérence de l'équation (Effectif du mois 1 + Entrées – Sorties = Effectif du mois 3). Si l'équation n'est pas valide, la réponse est supprimée. L'établissement sera corrigé de la non réponse totale.

3.1.2. Correction de la non-réponse totale par repondération

Sur les 100 000 établissements assujettis aux MMO (EMMO et DMMO), représentatifs de l'ensemble des établissements de 1 salarié ou plus sur le champ MMO, 85 000 répondent en moyenne chaque trimestre depuis 2011. Les établissements concernés par le processus de correction de la non-réponse sont :

- les établissements n'ayant pas répondu à l'EMMO,
- les établissements de l'EMMO/DMMO ayant répondu avec d'importantes incohérences (dans ce cas leurs réponses sont supprimées).
- les établissements n'ayant répondu à aucun des trois mois de la DMMO,
- les établissements de la DMMO ayant répondu partiellement mais n'ayant pu être imputés.

Taux de réponses exploitables trimestriel aux MMO (en %)

Année	Trimestre 1	Trimestre 2	Trimestre 3	Trimestre 4
2007	75,1	76,9	76,1	74,9
2008	76,9	76,7	74,7	74,7
2009	74,4	73,6	73,4	72,3
2010	73,2	71,9	70,9	65,5
2011	85,6	84,1	82,4	82,5
2012	87,7	85,8	84,2	83,8
2013	87,9	87,3	85,9	85,8

Note de lecture : au 4^{ème} trimestre 2012, 83,8 % des établissements assujettis et dans le champ des MMO renvoient une réponse exploitable.

La non-réponse totale est traitée par la méthode du ratio séparé. Pour cela, on post-stratifie l'échantillon par secteur (A38), tranche de taille (1 à 9 salariés, 10 à 49 salariés, 50 à 99 salariés, 100 à 249 salariés, 250 salariés ou plus) et zone géographique (Ile-de-France ou province). Au sein de chaque post-strate, le total d'une variable auxiliaire est connu ; il s'agit du nombre de salariés (variable présente dans la base de sondage).

La post-stratification permet de redresser les poids (pour des strates de taille suffisante) afin d'obtenir des estimations plus robustes. Cette méthode est destinée aux estimations portant sur le champ de la France métropolitaine. Elle consiste à dilater les poids initiaux (poids d'échantillonnage) des établissements répondants, afin de caler le nombre de salariés par strate sur des valeurs connues issues de la base de sondage (méthode de redressement par le ratio séparé).

Plus précisément, les nouveaux poids sont obtenus en multipliant le poids d'échantillonnage de l'établissement par un ratio calculé pour chaque post-strate. Le coefficient de redressement de la post-strate h est calculé de la façon suivante :

$$C_h = \left(\frac{\sum_{répondants \in h} d_i \cdot EFFREF_i + \sum_{non-répondants \in h} d_i \cdot EFFREF_i}{\sum_{répondants \in h} d_i \cdot EFFREF_i} \right)$$

avec

d_i : Le poids d'échantillonnage de l'établissement i ;

y_i : Le nombre de mouvement de type y de l'établissement i ;

$EFFREF_i$: L'effectif de référence de l'établissement i (effectif connu au moment du tirage).

Nous faisons ici l'hypothèse que la quantité suivante représente le nombre de salariés pour la post-strate h :

$$\sum_{répondants \in h} d_i \cdot EFFREF_i + \sum_{non-répondants \in h} d_i \cdot EFFREF_i$$

L'estimateur HT initial (avant redressement) du nombre de salariés pour la post-strate h est :

$$\sum_{répondants \in h} d_i \cdot EFFREF_i$$

Le poids du répondant i de la post-strate h redressé de la non-réponse est :

$$\hat{d}_i = d_i \cdot C_h$$

Cette méthode permet de caler l'estimateur HT avec les poids corrigés sur le nombre de salariés (EFFREF) par post-strate.

Remarque :

- a) La méthode décrite précédemment permet à la fois de corriger la non-réponse et d'effectuer un redressement par le ratio avec le nombre de salariés comme variable d'intérêt. Est-ce pour dire que si le nombre de salariés est bien corrélé aux mouvements de main d'œuvre, la précision est améliorée par ce calage ?
- b) Pour les MMO dans les DOM les post-strates de repondération sont définies par le secteur (A4), la tranche de taille (1 à 9 salariés, 10 à 49, 50 salariés ou plus) et le département ;
- c) Certains établissements renvoient une unique réponse pour un ensemble d'établissements couvrant parfois plusieurs secteurs, tranches de taille ou région. Ces établissements ne participent pas au calcul du coefficient de redressement de la non-réponse puisqu'il n'est pas possible de définir clairement leur post-strate.

3.2. Méthodes envisagées en DSN

Avec la généralisation de la DSN, la Dares envisage d'établir ses statistiques à partir de l'ensemble des DSN, donc de façon exhaustive sur l'ensemble du champ, et plus seulement sur le sous-champ des établissements de 50 salariés ou plus. L'amélioration de la précision qui en découlera permettra d'envisager des publications à des niveaux sectoriel ou géographique plus fins. En outre, cela permettra de prendre en compte les mouvements des établissements créés en cours d'année. Auparavant, les nouveaux établissements n'étaient ajoutés à l'échantillon qu'au moment du lancement de la campagne de collecte, une fois par an.

Actuellement, la DARES publie trimestriellement et annuellement des indicateurs statistiques en taux. Il s'agit, par exemple, de taux d'entrée (rapport entre le nombre d'entrées et le nombre de salariés sur une période) ou taux de sortie (rapport entre le nombre de sorties et le nombre de salariés sur une période). Ces indicateurs peuvent être déclinés par secteurs d'activité, tailles d'entreprise... À terme, le caractère exhaustif de la DSN devrait permettre de produire des statistiques en niveaux (niveaux des embauches, des licenciements...). Toutefois, le premier objectif que se fixe la Dares est de reproduire l'existant.

Dans tous les cas (niveaux, ratios, évolutions) le traitement de la non-réponse doit être étudié pour éviter des biais dans les statistiques de mouvements de main d'oeuvre. Les établissements non répondants peuvent en effet être justement ceux dont les caractéristiques où les évolutions sont particulières.

Il faut donc repérer les DSN manquantes. Pour cela, la difficulté majeure consiste à établir la liste de tous les établissements devant répondre à la DSN pour un mois donné. Théoriquement, tous les établissements employeurs doivent effectuer une DSN. Cependant, en DSN, la seule façon de repérer précisément si un établissement est employeur sera de constater qu'il a bien effectué une DSN (en DMMO on dispose actuellement de sources extérieures, comme les BRC ou les DADS, lesquelles seront substituées, comme la DMMO, par la DSN).

En effet, les répertoires d'entreprise ne contiennent pas une information suffisamment à jour sur la qualité d'employeur ou non des établissements. En pratique, la DSN étant obligatoire et contrôlée par les Urssaf avec des pénalités associées en cas de DSN absente ou en retard, la non-réponse devrait être limitée. Toutefois il paraît indispensable de la traiter pour produire des statistiques robustes à relativement court terme, en limitant l'impact sur ces statistiques des déclarations déposées en retard ou non parvenues à la DARES pour des raisons techniques (il faut en effet garder à l'esprit que les DSN sont déposées sur le site internet « net-entreprises.fr », contrôlées par l'Accoss, stockées par la

CNAV, transmises à la DARES, intégrées dans un premier SI de collecte puis ajoutées à la base des contrats avant l'établissement des MMO. À chaque étape il n'est pas à exclure qu'un incident se produise et qu'il faille enclencher les procédures de réémission, ce qui peut créer des retards ponctuels).

Une fois écoulé un délai au-delà duquel toutes les données relatives à un mois donné ont bien été reçues par le système DSN (ce délai ou recul sera défini empiriquement après expertise des fichiers reçus, mais on peut anticiper qu'il sera de l'ordre de 12 ou 18 mois), la base des contrats devrait pouvoir être considérée comme exhaustive. Les développements ci-dessous décrivent donc les méthodes actuellement envisagées par la DARES pour effectuer les redressements de court terme destinés à la production de statistiques conjoncturelles, sachant qu'une procédure de calage annuel sera également mise en place.

Il est essentiel de distinguer deux types de DSN manquantes :

- Un établissement a transmis sa DSN pour le dernier mois mais ne l'a pas fait pour un ou plusieurs mois antérieurs pour lesquels cette transmission était attendue (« trous de collecte », situation 1),
- Un établissement n'a pas transmis sa DSN pour le ou les derniers mois mais l'a fait pour les mois antérieurs (situation 2).

Ces deux situations ne seront pas traitées de la même façon. Dans le premier cas, la non-réponse est certaine du fait des règles propres à la DSN, car l'établissement aurait dû faire une DSN pour les mois concernés. Dans le second cas, il est en revanche pratiquement impossible de savoir à court terme, si l'établissement est non-répondant ou a cessé son activité. Par conséquent, dans le premier cas, la non-réponse d'un établissement sera traitée par une méthode d'imputation, alors que dans le second elle pourra être corrigée par l'une des deux méthodes décrites dans les paragraphes suivants (calage, « présents / présents »).

Situation 1 :

Dans cette situation, l'établissement a transmis une DSN portant sur le dernier mois (mois M) mais n'en a pas transmis pour un ou plusieurs mois antérieurs (mois M-1, M-2, ...). Les mois manquants pourront être imputés par une méthode similaire à celle décrite dans le paragraphe 3.1.1. Par exception, dans le cas particulier où l'établissement a un effectif nul à la fin du mois précédent le mois de non-réponse, une valeur nulle sera imputée. Dans les autres cas, la méthode de correction de non-réponse par imputation semble ici la plus appropriée. En effet, l'imputation (à la différence d'une méthode de calage) ne requiert pas d'information sur l'ensemble des établissements devant répondre à la DSN sur un mois donné. Or, lorsqu'on cherche à redresser les trous de collecte pour un mois donné, on ne peut être certains qu'aucun autre établissement n'a été en situation de « dernier » mois manquants à la même date. Par exemple, si l'établissement A a répondu en janvier et mars et l'établissement B seulement en janvier, alors on peut être certain que pour février la réponse de A était attendue mais dans le cas de B on ne peut jamais totalement exclure que son activité a cessé. Ainsi, une méthode de correction de la non-réponse par calage est délicate à appliquer car on n'est pas certain de la liste des établissements devant répondre en février. Pour corriger le trou de collecte de l'établissement A, il semble donc préférable de recourir à une méthode d'imputation qui ne nécessite pas de connaître le statut (répondant, non répondant ou cessé) de tous les établissements

Situation 2 :

Il s'agit ici de corriger la non-réponse d'un établissement n'ayant pas transmis sa déclaration sur le ou les derniers mois mais l'ayant fait sur les mois antérieurs. Deux méthodes sont pour l'instant envisagées par la Dares.

1^{ère} méthode :

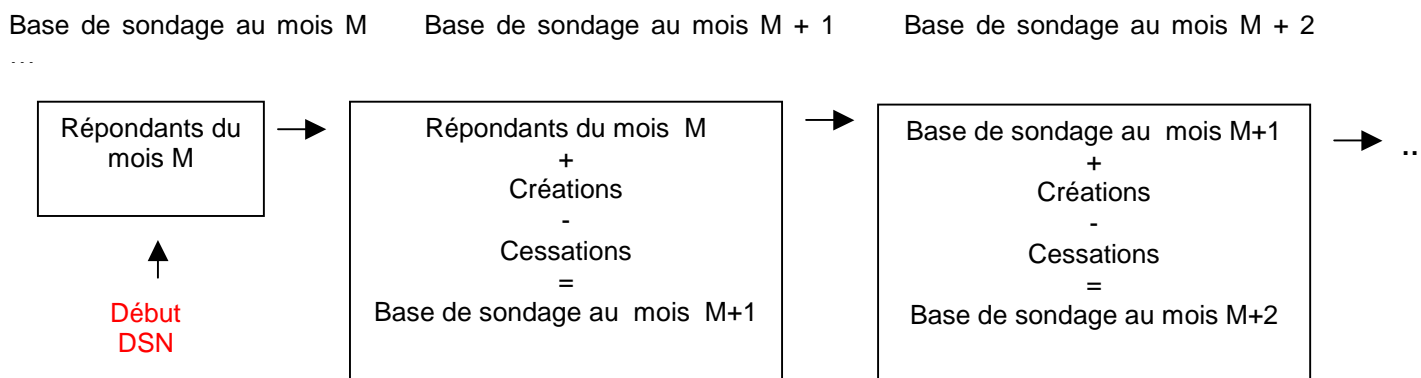
Le principe de cette méthode est de créer mensuellement une liste des établissements devant transmettre une DSN (base de sondage) puis d'appliquer une correction de la non-réponse totale « classique » par calage.

Création d'une base de sondage au mois M

Cette base de sondage reconstituée, pour le mois M, sera un agrégat des :

- Etablissements ayant effectué une DSN au mois M-1 et non cessés,
- Nouveaux établissements créés (repérés par les nouveaux établissements effectuant une DSN).

Evolution de la base de sondage



Cette méthode ne permet toutefois pas de détecter les nouveaux établissements devant répondre mais ne répondant pas. De plus, elle nécessite d'actualiser efficacement la liste des établissements cessés. Un établissement cessé ne doit pas être redressé de la non-réponse mais doit être supprimé de la base de sondage. Or, l'information sur la cessation n'est pas toujours à jour dans les répertoires d'entreprise ou est déclarée avec retard. Une partie du travail de gestion des DSN à la Dares sera donc de s'assurer que les établissements non répondants et pour lesquels les dernières données disponibles suggèrent qu'ils ont encore des contrats en cours, sont en fait cessés. Globalement, la base de sondage sera tout de même imparfaite à court terme ce qui constitue le principal inconvénient de cette méthode.

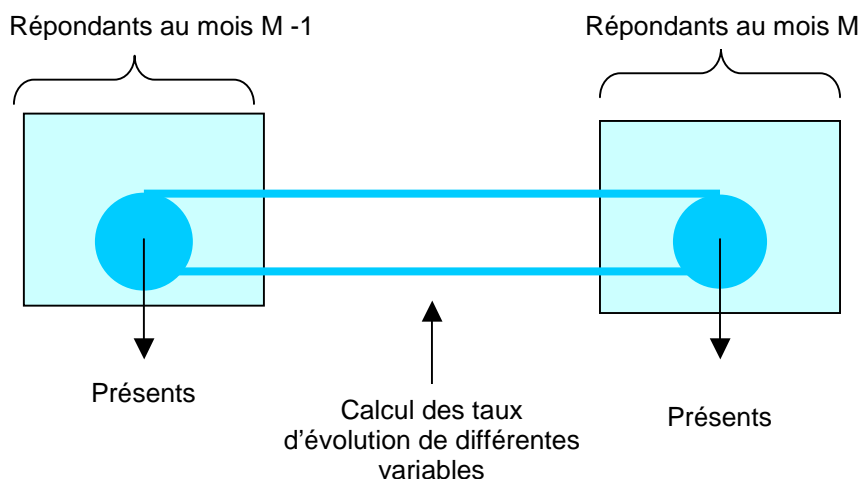
Correction de la non-réponse totale

Une fois la base de sondage reconstituée pour un mois donné, la DARES procèdera à la correction de la non-réponse totale. En théorie, chaque établissement de cette base de sondage doit transmettre une DSN. Un poids de 1 sera donc attribué à chaque établissement. Après analyse du processus de non réponse ce poids sera corrigé. Diverses techniques de calage pourront être envisagées. La méthode actuelle du redressement de la non réponse totale (ratio séparé sur la variable dernier effectif connu) pourrait être à nouveau utilisée (voir paragraphe 3.1.2).

2ème méthode :

Le principe de cette méthode est d'établir des statistiques en évolution sur les établissements « présents / présents » et d'appliquer cette évolution au dernier point « exhaustif » connu.

Chaque mois, des établissements transmettent leur DSN. Les taux d'évolution pourront être calculés entre deux mois consécutifs. Les établissements pris en compte seront les établissements répondant au mois M et M-1.



Cette méthode nécessite un point d'ancrage à mettre à jour régulièrement pour prendre en compte la démographie des établissements. Les taux d'évolution calculés seront appliqués sur les stocks de la période de référence (point d'ancrage). Ce procédé permettra de calculer des statistiques en taux mais également en niveau.

L'établissement d'un point d'ancrage est une étape délicate de cette méthode. En effet, comme pour la 1^{ère} méthode, il est nécessaire de connaître les créations ainsi que les cessations d'établissements.

Exemple : Calcul du taux d'entrée en CDD, exploitation du mois M

Étape 1 : Calcul du taux d'entrée en CDD du mois M (T_M) sur les « présents » / « présents » M-1/M. Calcul du taux d'entrée en CDD du mois M-1 (T_{M-1}) sur les « présents » / « présents » M-1/M.

Étape 2 : Calcul du taux d'évolution entre T_M et T_{M-1} .

Étape 3 : Calcul du taux d'entrée en CDD du mois M-1 (T^*_{M-1}) sur l'ensemble des établissements

Étape 4 : On applique le taux d'évolution calculé à l'étape 3 à T^*_{M-1} . On obtient une estimation du taux d'entrée en CDD au mois M.

Le principal inconvénient de cette méthode est qu'elle ne permet pas de prendre en compte la démographie des établissements à court terme car les établissements créés et cessés ne sont pas présents/présents. En revanche, elle ne nécessite pas de maintenir et de mettre à jour une base de sondage. Son implémentation est donc plus aisée.

Conclusion

Cet article présente, à travers la discussion des choix méthodologiques effectués (ou en cours d'élaboration) par la Dares pour reconstituer les mouvements de main-d'œuvre à partir de la nouvelle source administrative qu'est la DSN, les principales particularités de ces données mensuelles issues de la paie, dont il faut tenir compte pour établir des statistiques d'emploi. Il en ressort deux messages principaux. En premier lieu, des retraitements plus ou moins complexes sont nécessaires pour affecter les éléments liés à la rémunération aux périodes d'activité réelle auxquelles elles se rapportent. En second lieu, l'exploitation statistique de ces données nécessitera de recueillir de l'information sur les établissements, notamment sur les cessations et changement de Siret. Or, cette information est aujourd'hui incomplète ou disponible avec retard dans les référentiels d'établissement existants (comme Sirene).

Malgré ces difficultés, la DSN est potentiellement une source très riche pour l'analyse de l'emploi : exhaustive et rapidement disponible, elle devrait permettre d'étendre le champ des exploitations statistiques à la fois en matière d'analyse conjoncture et d'études plus structurelles. S'agissant des mouvements de main-d'œuvre, elle devrait permettre de publier plus fréquemment, sur des données plus fraîches, et de produire une information plus détaillée. En effet, au-delà de la précision apportée par l'exhaustivité de la source, la DSN permettra également de décomposer plus finement les embauches par type de contrat (en distinguant par exemple les contrats aidés) et les sorties par motif de rupture (qui sont plus détaillés en DSN qu'en MMO).

Il convient enfin de rappeler que la DSN est une source en gestation, et que l'organisation du système d'information dont elle est issue est complexe. La raison en est paradoxalement l'objectif même de simplification des formalités administratives des entreprises qui a donné naissance à ce projet : la coordination des besoins multiples et hétérogènes des organismes utilisateurs (en premier lieu pour la protection sociale, mais aussi pour les statistiques) et la contrainte de ne pas demander aux entreprises une information qui peut être reconstituée à partir des données natives, impose à la fois une normalisation stricte et stable des données, qui permet seule de donner un cadre cohérent aux projets de reconstitution menés par ces organismes, et une communication précise auprès des parties prenantes du projet (les entreprises mais aussi les éditeurs de logiciels de paie). Pour toutes ces raisons, au-delà de l'exhaustivité des déclarations, leur qualité – au sens de la conformité à la norme – constitue un enjeu important pour celle des statistiques. On ne dispose pas encore d'un recul suffisant pour évaluer cet aspect.