

Reconstitution des mouvements de main-d'œuvre à partir de la déclaration sociale nominative

Quelle correction apportée dans la phase de montée en charge du dispositif ?

Journées de Méthodologie Statistique
Juin 2018

Kévin Milin
Dares

Les mouvements de main-d'œuvre

Mesurent **les flux de main-d'œuvre** selon :

- Le type de mouvement ;
 - Les **entrées** (*i.e.* embauches en CDD ou CDI) ;
 - Les **sorties** (les fins de CDD, ou ruptures de contrat par motif) ;
 - Les entrées et les sorties ne sont pas reliées.
- Les caractéristiques des salariés (âge, sexe, catégorie socio-professionnelle, ...) ;
- Les caractéristiques des établissements (secteur d'activité, nombre de salariés, ...) ;
- Le nombre de mouvements est traditionnellement rapporté à l'effectif

Changement de source pour les MMO

Les mouvements de main-d'œuvre (MMO) étaient construits à partir :

- D'une déclaration administrative (mensuelle) pour les établissements de plus de 50 salariés (**DMMO**) ;
- D'une enquête (trimestrielle) sur les moins de 50 salariés (**EMMO**).

Changement de source pour les MMO :

- Introduction des déclarations sociales nominatives (**DSN**) dès 2013
 - Objectif : se substituer à un grand nombre de déclarations administratives, dont la DMMO
- Publication des données MMO suspendue au 1^e trimestre 2015
 - Trop peu de répondants, en lien avec la montée en charge des DSN

Reconstitution des MMO à partir des DSN

- DMMO/EMMO : variables de flux (déclaration des mouvements de main-d'œuvre)
- DSN : stock de contrats (reliés aux logiciels des payes du mois)
- Reconstruire les MMO, c'est comparer la liste des contrats en cours sur deux mois successifs pour identifier les embauches et les fins de contrats
- De nombreux apports avec les DSN :
 - Entrées et sorties sont désormais reliées
 - Ajout de dimensions d'analyses (motifs de recours notamment)
 - Capacité de calcul des effectifs et des mouvements dans toutes les dimensions (auparavant, seul l'effectif total de fin de trimestre/mois était déclaré)
 - Dimensions géographiques et sectorielles plus fines

Construire une méthodologie permettant le redressement des données DSN

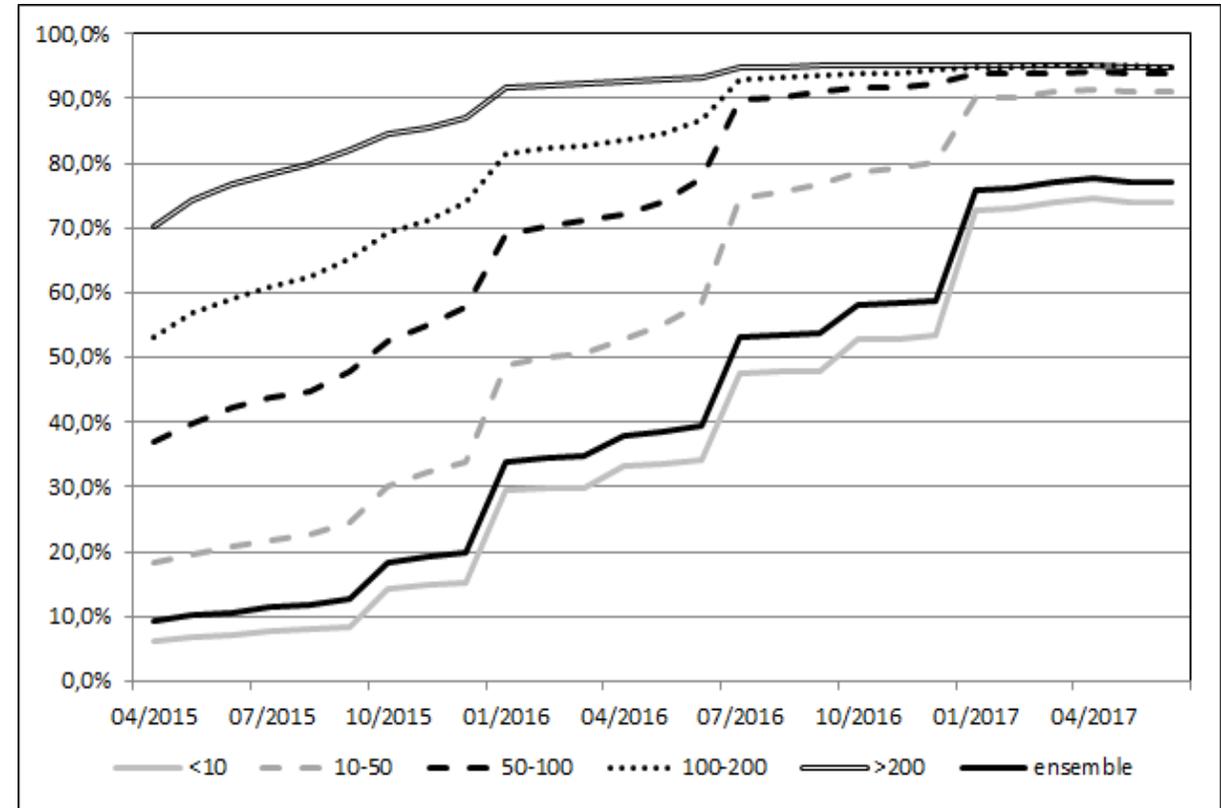
Premières étapes :

- Etape 1 : construire des référentiels correspondant au champ « MMO »
 - prise en compte partielle de la démographie des entreprises

- **Etape 2 : s'assurer que les répondants DSN sont représentatifs de l'ensemble du champ (i.e. correction de non-réponse et calage)**

Car:

- Il y a de la non-réponse (et c'est normal)
- Différents paliers sont identifiables, ce qui implique différentes variantes temporelles des modèles de redressement



Taux de réponse à la DSN sur le champ MMO

Principe de la méthodologie de redressement de la non-déclaration

Différentes étapes :

a. **Estimation de la probabilité de déclaration des établissements**

- Secteur d'activité, région d'activité, effectif et date de création de l'établissement
- Chiffre d'affaire, catégorie juridique et nombre d'établissements de l'entreprise
- Variables d'interaction
- **Variables quantitatives catégorisées pour prendre en compte les non-linéarités**

b. Création de classes de pondération par la méthode d'Haziza Beaumont (2007)

c. Pondération = inverse du taux de réponse dans chaque classe

d. Calages sur marges (effectifs)

- Différentes dimensions : secteur d'activité, taille de l'établissement, zone géographique

Focus sur l'estimation de la probabilité de déclaration

- Modèle (logit) estimé chaque mois pour chaque grand secteur
- ... mais **une spécification différente par palier**
 1. Découpage des distributions de l'effectif de l'établissement et du chiffre d'affaire de l'entreprise par le moyen d'un jeu de 20 quantiles.
 2. Estimation du modèle complet (sauf les variables d'interactions composées des variables à catégoriser).
 3. Application des tests d'égalité suivants :
 $\alpha_i = \alpha_{i+1}$ et $\beta_j = \beta_{j+1}$ avec α_i (resp. β_j) le paramètre associé à la i -ème (resp. j -ième) classe d'effectif (resp. du chiffre d'affaire)
 4. Fusion de la j -ième et $(j+1)$ -ième classe d'effectif en lien avec le test qui présente la statistique de test la plus proche de 0 (idem pour les classes du chiffre d'affaire) ;
 5. Répétition des étapes 2, 3 et 4 jusqu'à ce que tous les tests d'égalité soient rejetés au niveau 5 %.

Prise en compte des non-linéarités : illustrations

- Plus le chiffre d'affaire est important, plus le coefficient de la classe associée est fort. Le coefficient de la dernière classe est particulièrement élevé.

chiffre d'affaire en k€	?	< 0	0-56	56-91	91-118	118-143	143-193	193-220	220-383	383-1566	1566-7802	7804 et +
coefficient	0	-1,2	-1,4	-0,4	0,0	0,1	0,2	0,3	0,4	0,4	0,5	1,3

- La classe d'effectif avec le coefficient estimé le plus fort, n'est pas celle qui est constituée des plus grands effectifs.

effectif	1	2	3-8	9-15	16-22	23-47	48 et +
coefficient	-1,0	0,1	0,2	0,3	0,2	0,1	0,0

Quelques éléments rassurants sur la qualité des modèles : des statistiques non biaisées (1)

- Démarche :
 - Le modèle de redressement explique une partie de la non-réponse, mais ne peut pas tout expliquer (aucun modèle n'est parfait) ;
 - Il n'y aura pas de biais dans les statistiques finales si la partie qui n'est pas expliquée n'est pas corrélée avec ces statistiques finales ;
 - Pour le tester, on essaie de diminuer la part d'inexpliqué du modèle grâce à une variable proche de nos statistiques finales et disponible sur tout le champ.
- Variable auxiliaire utile : déclaration préalable à l'embauche (DPAE)
 - Très corrélée aux MMO ;
 - Connue de manière exhaustive sur tout le champ.

Quelques éléments rassurants sur la qualité des modèles (2) : des statistiques non biaisées

- Pour chaque mois, on considère le modèle de réponse, et on ajoute la variable DPAE correspondante
 - Idée : si le mécanisme de non-réponse est ignorable, la DPAE n'explique pas la non-déclaration
 - On teste donc la nullité des coefficients associés

	Construction	Industrie	Tertiaire	Ensemble
Taux d'entrée en CDD	75 %	70 %	70 %	72 %
Taux d'entrée en CDI	85 %	95 %	70 %	83 %

- Part des mois entre mai 2015 et décembre 2016 pour lesquels la partie non-expliquée des modèles n'est pas corrélée avec les DPAE
- **Conclusion : il ne semble pas y avoir de biais lié à la non-déclaration dans les statistiques finales**

Quelques éléments rassurants sur la qualité des modèles (3) : évaluation de l'erreur totale

- Démarche :
 - Recherche d'une variable connue sur toute la population concernée et corrélée avec les mouvements de main-d'œuvre ;
 - Calcul de la « vraie » valeur de cette variable sur tout le champ
 - Restriction du champ aux seuls déclarants DSN et estimation de cette « vraie » valeur au moyen des modèles de redressement de la non-réponse ;
 - Taille de l'erreur = écart entre la vraie valeur et la valeur estimée ;
- Variable utilisée : les DPAE (très corrélées avec nos statistiques de mouvements de main-d'œuvre)

Quelques éléments rassurants sur la qualité des modèles (4) : évaluation de l'erreur totale

• Conclusions :

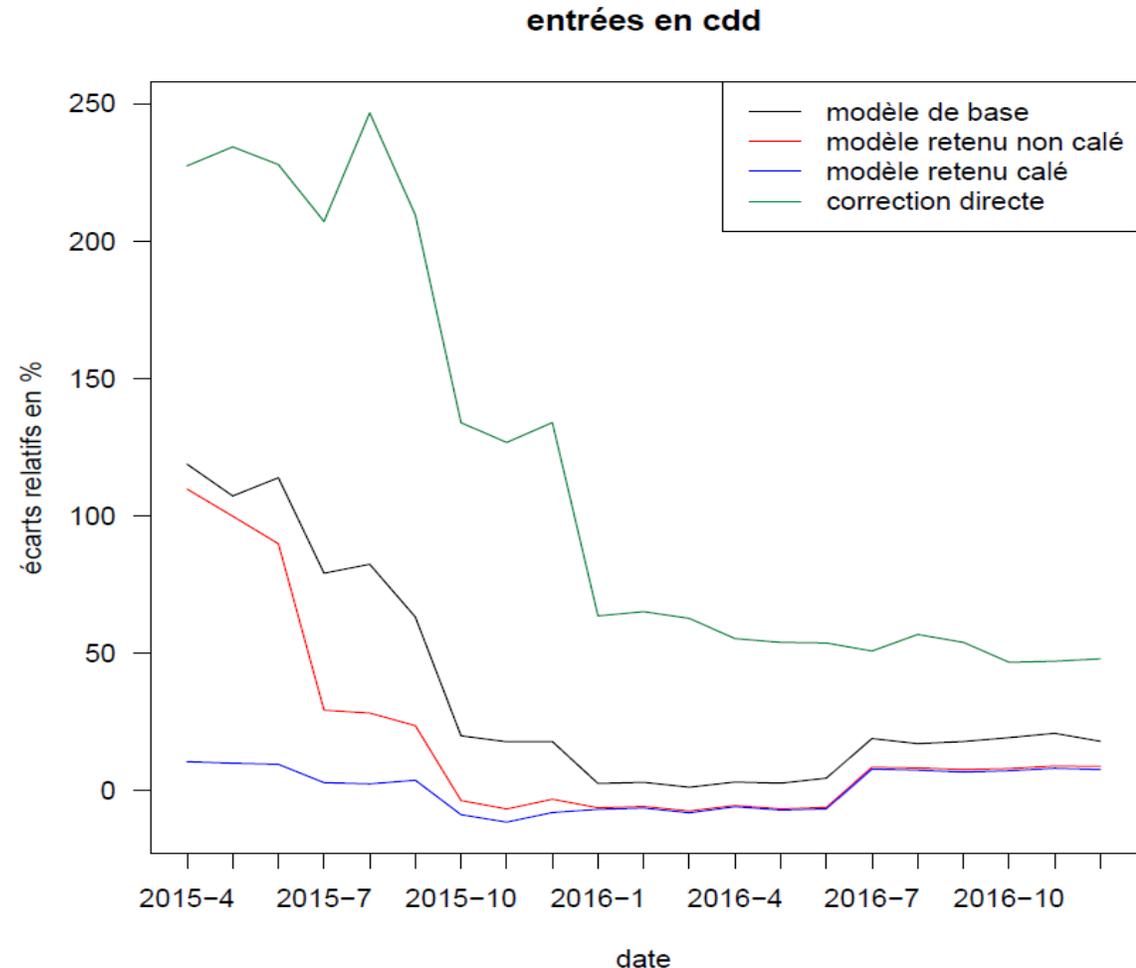
- Le choix de mettre en place une correction de non-réponse est confirmé par ces évaluations

- (comparaisons entre la courbe rouge et la courbe verte) [*i.e.* les déclarants DSN ont donc plus de mouvements les autres]

- La qualité du redressement s'améliore significativement à partir du 3^e trimestre 2015 ; l'effet du calage sur marges est important sur 2015.

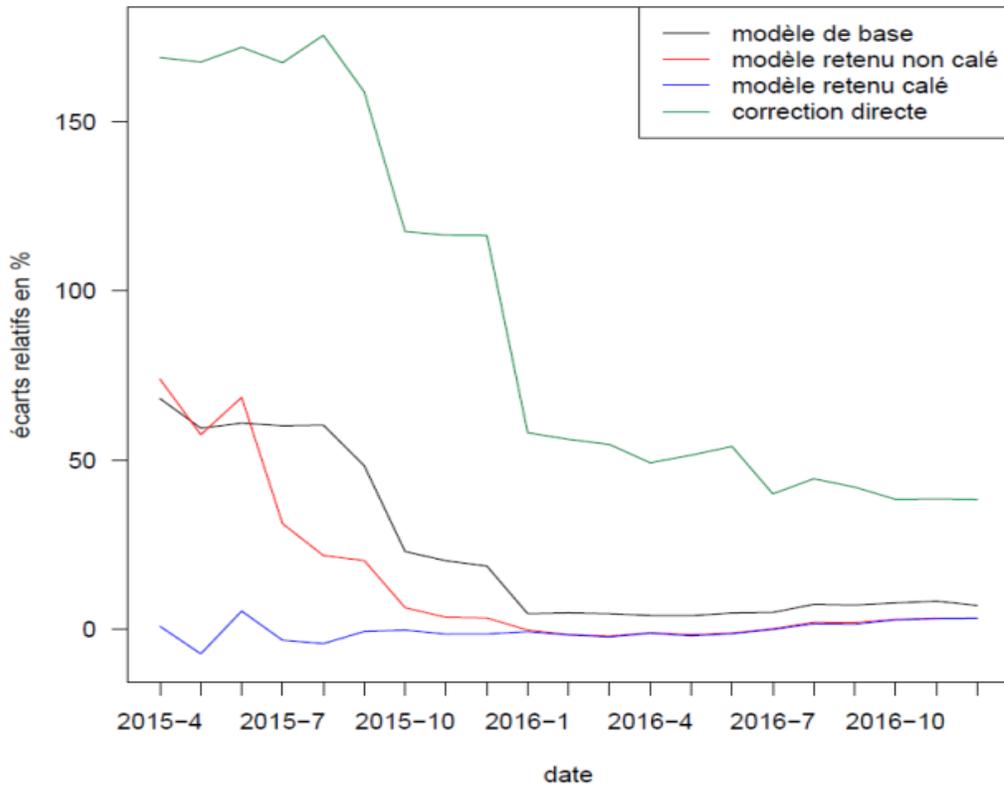
- Importance de la prise en compte des non-linéarités dans la construction des probabilités de déclaration

- Sans cela, les entrées sont surestimées (courbe noire)

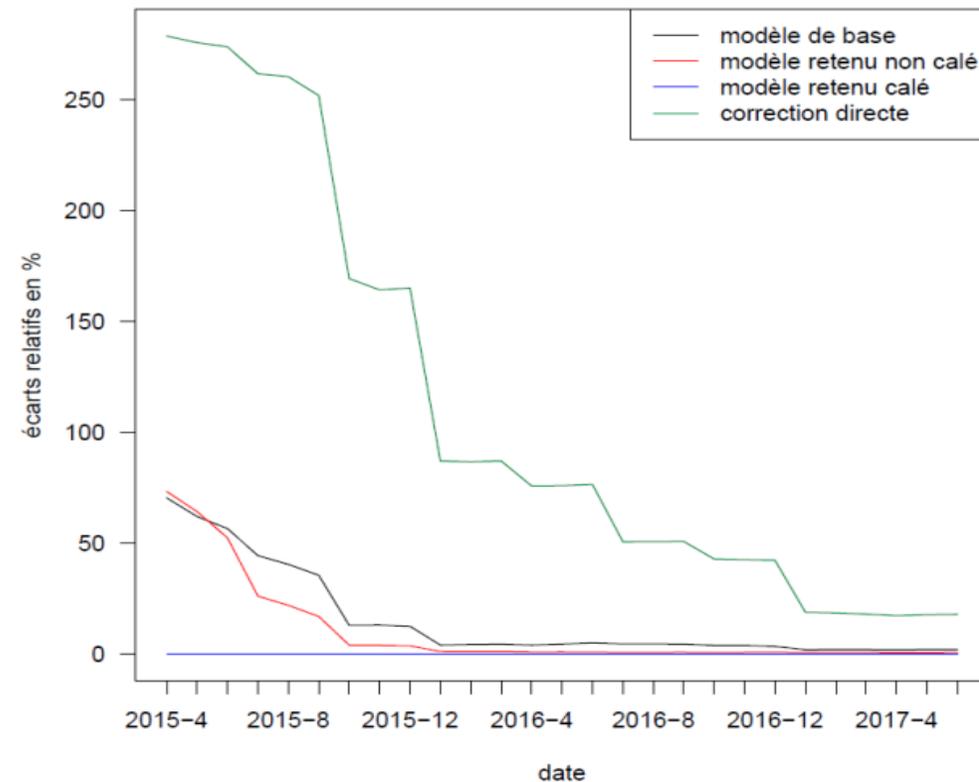


Quelques éléments rassurants sur la qualité des modèles (5) : évaluation de l'erreur totale

entrées en cdi



effectif



Deux autres étapes pour aboutir aux séries finales

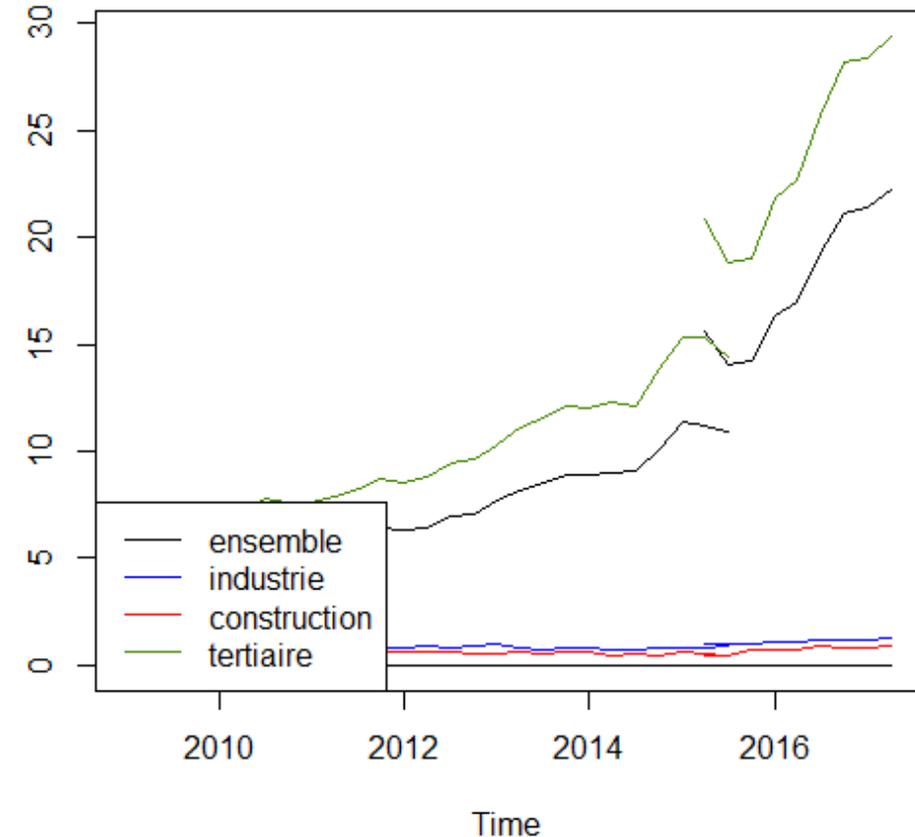
- Etape 3 : corriger les problèmes de qualité dans les déclarations DSN
- Etape 4 : corriger les données historiques des biais de sous-déclaration (rétropolation au niveau des données détaillées)

→ L'ensemble de la méthodologie est décrite dans un document d'étude, publié le 21 juin prochain

Accompagné

➤ D'une reprise de la diffusion des séries

➤ D'une étude sur les évolutions des embauches/ruptures de CDD/CDI ces 25 dernières années



Taux de fin de CDD très courts, par secteur, avant rétropolation

Merci de votre attention