

Groupe de travail du Cnis

« Mesurer les discriminations : défi statistique et méthodologique »

Compte rendu de la 3^e réunion - 13/05/2022

Président : Michel Villac (Haut Conseil de la famille, de l'enfance et de l'âge)

Rapporteurs : Sylvie Le Minez (Insee), Clément Martin (Défenseur des droits), Patrick Simon (Ined)

Destinataires : les membres du groupe de travail

Ordre du jour : Présentation de 2 interventions, du projet de plan du rapport et du programme prévisionnel des prochaines séances

1. « L'étude des discriminations dans les travaux impulsés par la Dares » Emilie Arnoux, Pierre Villedieu (Dares)
2. « Inégalités selon l'origine : données et approches économétriques », Roland Rathelot (Institut Polytechnique de Paris)
3. Projet de plan de rapport (Patrick Simon)
4. Programme prévisionnel des prochaines séances (Sylvie Le Minez)

Participants

Pour les représentants du GT : Michel Villac (président), Martin Clément (rapporteur), Patrick Simon (rapporteur), Sylvie Le Minez (rapporteur), François Guillaumat-Tailliet (CNIS), Arnaud Montus (CNIS)

Pour les membres du GT :

Jean-Francois Amadiou, Université Paris 1

Franck Arnaud, Drees

Emilie Arnoult, Dares

Nathalie Bajos, Inserm

Didier Bourgoin, FSU

Eric Briat, ANCT

Marielle Chappuis, Défenseur des droits

Maryama Diakhaté, SSM Justice

Sophie Elizeon, Dilcrah

Hélène Garner, France stratégie

Stéphane Jugnot, CGT

Thomas Le Barbanchon, Université Boccoli, Crest

Yannick L'Horty, Université Gustave Eiffel et TEPP-CNRS

Cédric Lorant, Unanimes

Yara Makdessi, SSM Justice

Camila Marin, Unanimes

Fabrice Murat, Depp

Jean-Luc Outin, Conseil national des politiques de lutte contre la pauvreté et l'exclusion sociale

Charlotte Parez, Medef

Roland Rathelot, IPPAris, Ensae, Crest

Jean-François Ravaud, Inserm

Christian Rieg, DGAFP

Mirna Safi, Sciences Po, OSC, LIEPP & Crest-LSQ

Jan Robert Suesser, LDH
Cédric Schönwald, Solution Scribe (sous-titrage)
Gabriel Thoison, FO
Chantal Vesserau, Pôle emploi
Delphine Vessillier, CPME
Augustin Vicard, SSM Jeunesse et sports
Pierre Villedieu, Dares, Sciences-Po, LIEPP

1. « L'étude des discriminations dans les travaux impulsés par la Dares » Emilie Arnoux, Pierre Villedieu (Dares)

Ont tout d'abord été présentées les mesures des discriminations par voie de testings réalisées à la Dares, avec quatre applications en particulier :

- [un testing auprès de quarante grandes entreprises en 2016](#) pour mesurer la discrimination à l'embauche selon l'origine, avec 30 à 40 offres testées par entreprises;
- [un testing sur les décrocheurs scolaires en 2018](#), avec 5000 offres d'emploi testées et 10 000 candidatures spontanées pour les métiers de cuisinier et de maçon ;
- [un testing « contrats aidés » en 2018](#) pour mesurer le retour à l'emploi de femmes éloignées de l'emploi selon qu'elles sont ou non récemment passées par un contrat aidé avec 3000 offres testées pour les métiers de secrétaires et aides-soignantes.
- Puis le [testing de grande ampleur](#) réalisé plus récemment (entre décembre 2019 et mars 2020, en juillet 2020, en avril 2021), sur un peu plus de 2 200 entreprises de toutes tailles, dans l'ensemble des départements (et plus de 1 100 communes, dont 23 % en Île-de-France) avec 12 métiers testés (pour faire varier les niveaux de qualification, les métiers en tension ou pas, le degré de féminisation des métiers). Ce testing a été réalisé par ISP-Corum et l'IPP pour la Dares et répondait à une commande publique. L'objectif était de tester la discrimination générale à l'embauche selon le sexe sur différentes dimensions (origine, âge, niveau de qualification...).

La Dares a ensuite présenté l'outil Namsor, un outil d'inférence de l'origine par les noms/prénoms, qui peut notamment être mobilisé pour étudier des écarts de situation/discriminations éventuelles : cet outil permet d'assigner un sexe ou un pays d'origine à une personne à partir de son nom et de son prénom. Cet outil est utilisé dans des travaux de recherche, par des entreprises, par la [Commission européenne](#), et à la Mission d'animation de la recherche de la Dares. Cet outil évolue avec le corpus d'entraînement et ses prédictions peuvent donc évoluer avec le temps. Les requêtes se font *via* une API. La Dares a montré un exemple d'utilisation de cet outil probabiliste mobilisant des données administratives (dans lesquelles l'origine géographique n'est pas connue).

Enfin la Dares a présenté un projet d'expérimentation d'évaluation d'une mesure de lutte contre la discrimination à l'embauche (tester l'impact d'une formation « recruter sans discriminer » auprès des recruteurs d'entreprises, avec un échantillon d'entreprises mettant en place cette formation et un échantillon témoin d'entreprises similaires ne mettant pas en place cette formation), né notamment des résultats du testing de 2016 (avec des entreprises conscientes des processus discriminatoires) et de la loi Égalité, Citoyenneté de 2017. Il a cependant été abandonné en raison de la crise sanitaire, mais donne à voir comment des mesures anti-discriminations peuvent être évaluées.

2. « Inégalités selon l'origine : données et approches économétriques », Roland Rathelot (Institut Polytechnique de Paris)

Roland Rathelot a présenté un ensemble de données (notamment d'origine administrative utilisées à des fins de statistique) et d'études portant sur les discriminations selon l'origine (nationalité, pays de naissance des personnes et/ou de leurs parents), axées principalement sur le marché du travail.

En préambule, il a rappelé qu'il était important de viser un système d'observation et des méthodes permettant de mesurer les écarts selon l'origine dans les différents « marchés », de comprendre les mécanismes conduisant à ces écarts (ce qui relève de la discrimination, mais aussi de l'origine socio-économique, du lieu de résidence ou des inégalités scolaires par exemple), de comprendre les causes de la discrimination quand discrimination il y a, et enfin d'évaluer l'impact de mesures visant à réduire les écarts de situation observés. Il a également rappelé l'importance de la population avec des origines migratoires en France et les écarts importants de taux d'emploi selon l'origine, même en contrôlant des caractéristiques individuelles, comme le niveau d'éducation des personnes et leur origine sociale (avec par exemple des taux d'emploi inférieurs de 7 à 8 points entre les descendants d'immigrés d'origine africaine et les personnes sans ascendance migratoire).

En France, si plusieurs enquêtes permettent de mesurer les écarts de situation sur le marché du travail selon l'origine (enquêtes Emploi, Génération et Trajectoires et origines), ce n'est pas vraiment le cas des données administratives (comme les bases tous salariés, très usitées dans le champ du marché du travail, reposant sur les DADS, ou la DSN aujourd'hui, et qui ne comprennent que la nationalité actuelle des personnes en 3 postes ; le fichier historique des demandeurs d'emploi, un peu plus riche sur la nationalité actuelle). A cela s'ajoutent le recensement de la population (qui permet d'identifier les immigrés et peut-être à l'avenir les personnes ayant des parents nés à l'étranger) ainsi que l'échantillon démographique permanent (appariement pour un échantillon de données du recensement et de sources administratives, mais vite limité en raison de sa taille). A l'étranger existent dans certains pays des données administratives plus riches, comme précisé dans la présentation de Roland Rathelot (avec de nombreux exemple de sources administratives, parfois exhaustives, contenant beaucoup plus d'informations sur les origines : en Suède (dispositifs LINDA et IFAU-LEED), au Danemark (IDA), en Autriche (ASSD et Register-base Labour Market Career), au Brésil (RAIS), en Allemagne (appariement de l'enquête SOEP avec des données administratives), au Portugal (Quadros de Pessoal)). Dans d'autres pays, on est plus proche de la situation française, comme en Allemagne (avec cependant des appariements de la grande enquête SOEP avec des données administratives sont répandus) ou encore au Portugal (Quadros de Pessoal).

Si les testings sont irremplaçables pour mesurer les écarts de taux de réponse à des candidatures sur le marché du travail, ils ne sont pas si faciles à interpréter en termes d'écarts de taux d'embauches, et il est très important d'avoir accès à un riche jeu de variables de contrôle (comme, par exemple, l'historique détaillé des épisodes de chômage et d'emploi des personnes pour comparer des individus comparables à l'exception de leurs origines, dont on dispose dans des sources administratives) et au-delà de l'embauche, il est important d'observer dans les entreprises les écarts de salaire, de promotion, d'accès à la formation, etc. et de les comprendre.

Les données administratives en France, comme à l'étranger, ont permis de faire des progrès substantiels dans la compréhension des écarts de situation sur le marché du travail selon le genre depuis plus d'une dizaine d'années. Roland Rathelot a cité plusieurs travaux de recherche menés en France et à l'étranger à partir de données administratives ou issues de plateformes (comme RBnB), portant sur des écarts Femmes-Hommes ou selon l'origine et permettant de comprendre certains mécanismes sous-jacents.

3. Projet de plan de rapport

Patrick Simon a esquissé la trame du rapport telle qu'envisagée par le président et les rapporteurs du GT, qui a donné lieu à un premier échange entre les membres.

Le rapport serait composé d'une introduction (définitions et approche des discriminations, enjeux de sa mesure, mandat et composition du GT), d'une première partie sur les méthodes de mesure des discriminations, d'une deuxième partie sur les motifs de discrimination (notamment leurs traductions en catégories statistiques), d'une troisième partie sur un état des lieux des données de la statistique publique par domaines (emploi, logement, éducation ...) et motif, d'une quatrième partie sur le cadre juridique et les politiques publiques, d'une cinquième partie sur les besoins des utilisateurs et des personnes rencontrées et enfin d'une sixième partie sur les scénarios d'évolution et les préconisations.

4. Programme prévisionnel des prochaines séances

Sylvie Le Minez a présenté le programme prévisionnel des prochaines séances envisagé par le président et les rapporteurs du GT, qui a été pensé en articulation avec le plan du rapport.

La séance prochaine du 23 juin sera consacrée aux besoins des utilisateurs, en s'appuyant pour commencer sur un recueil des besoins auprès de tous les membres du GT, qui sera complété à la rentrée de septembre et ouvert plus largement.

A la rentrée, deux séances devraient être consacrées aux motifs de discrimination et aux catégories de population discriminées (quelle traduction statistique des motifs de discrimination ? quelle couverture des populations concernées ? avec des zooms sur certains motifs : handicap, orientation sexuelle et de genre, discriminations syndicales, etc.). Deux autres séances devraient ensuite être consacrées à l'état des lieux des sources et à des approfondissements sur des méthodes (sources mobilisables par domaine, questionnement sur les discriminations vécues/ressenties, l'analyse de la ségrégation au moyen du recensement notamment). A partir de janvier 2023, des séances devraient être consacrées au cadre juridique, aux besoins/missions de la statistique publique et des autres acteurs, aux stéréotypes... et bien sûr à des échanges sur les préconisations du rapport jusque fin juin 2023.

La prochaine réunion aura lieu le 23 juin 2022.