



LA MESURE DES DISCRIMINATIONS LIÉES À “L’ORIGINE”

**Statistiques et testing présentés lors du colloque
du 22 octobre 2007 à Lyon**

colloque et publication sous la
direction scientifique d’Eric Cédiey



ISM CORUM
32 cours Lafayette – 69421 Lyon cedex 03 – France
+33 (0)4 7284 7890 / dir@ismcorum.org / www.ismcorum.org

Eric Cédiey (dir.), ISM-Corum (éd.), 2008

La mesure des discriminations liées à "l'origine"

Statistiques et testing présentés lors du colloque du 22 octobre 2007 à Lyon

téléchargement : www.ismcorum.org, Références, Thème : Lutte contre les discriminations,
Domaine : Diagnostics et Etudes, Date : Novembre 2008.

Sommaire

Propos du colloque 3

Glossaire - *Eric Cédiey, ISM CORUM, Benoît Riandey, INED* 5

CADRE PROBLEMATIQUE

Que mesure-t-on ? Pour quoi ? Comment ? - *Eric Cédiey, ISM CORUM* 11

**Deux logiques de mesure : enquêtes par sondage et représentativité,
ou enquêtes cas-témoins et test de causalité** - *Benoît Riandey, INED* 15

CADRE REGLEMENTAIRE

La CNIL et la mesure des discriminations liées à « l'origine »
Sophie Vulliet-Tavernier, directrice des affaires juridiques et de l'expertise, CNIL 21

ENQUETES PAR SONDAGE EN POPULATION GENERALE

**Enquête FQP et enquête Structure des Salaires –
Études sur la diversité dans la fonction publique et sur les différences salariales
dans le secteur privé** - *Julien Pouget, INSEE* 29

**L'enquête européenne TIES : The Integration of European Second Generation –
Méthode d'échantillonnage : patronymes et screening téléphonique** -
Christelle Hamel, INED 39

ANALYSES DE COHORTES

**Différenciations des parcours professionnels des jeunes
selon l'origine socioculturelle** - *Thomas Couppié, CEREQ* 49

**Méthodes et résultats de l'étude « Porter un prénom arabe ou musulman
est-il discriminant dans l'enseignement supérieur ? »** - *Marie-Noëlle Decharne
et Eric Liedts, Pôle Universitaire de Lille Nord-Pas de Calais* 57

ETUDES A L'INTERIEUR DES ORGANISATIONS

L'enquête expérimentale « Mesure de la diversité » : test de différentes catégorisations des origines et réaction des enquêtés - <i>Patrick Simon, INED</i>	71
« Testing sollicité » au sein du Groupe Casino - <i>Fabrice Foroni, ISM CORUM</i>	81
L'expérience de l'ethnic monitoring aux Pays-Bas. Recensement interne des origines des salariés par le lieu de naissance de leurs parents et objectif de rééquilibrage de leurs proportions dans le cadre de la loi Samen, 1998-2003 - <i>Virginie Guiraudon, CNRS</i>.....	93
Remarques finales <i>Sophie Vulliet-Tavernier, directrice des affaires juridiques et de l'expertise, CNIL</i>	99
Principales références en ligne	100

Propos du colloque

La mesure des discriminations liées à « l'origine » suscite de nombreux débats, scientifiques et politiques, centrés sur la question dite des « statistiques ethniques ». Comment produire des données permettant de mesurer les discriminations et d'évaluer les politiques mises en place ? Peut-on classer les personnes en fonction d'un critère d'origine « ethnique » ? Si oui, quelles définitions retenir pour caractériser cette « origine » et comment nommer les différents groupes concernés ?

Ces questions de classement et de catégories renvoient en fait à des questions de fond qui méritent d'être précisées dans les débats, et qui concernent notamment l'objet de la mesure, ses objectifs, et l'échelle pertinente de son application :

- Que se propose-t-on exactement de mesurer : la diversité des populations, la présence de pratiques discriminatoires, les inégalités de situation ?
- À quels objectifs et à quelles politiques, publiques ou privées, peuvent servir ces différents exercices de mesure : créer un référentiel de la diversité, corriger les pratiques discriminatoires, suivre l'évolution des écarts de situation pour évaluer les politiques ?
- À quelle échelle est-il pertinent de réaliser ces différentes mesures : sur l'ensemble de la population ? dans des espaces sociaux déterminés, comme les organisations de travail ? faut-il classer et représenter tout le monde ou s'appuyer sur des échantillons témoins ?

Le colloque du 22 octobre 2007 sur *La mesure des discriminations liées à "l'origine"* a voulu alimenter la réflexion autour de ces questions, à partir d'une présentation au public des différentes techniques de mesure actuellement utilisées, en particulier dans le champ de l'emploi : statistiques en population générale, suivis de cohortes, testings et testings sollicités, comparaisons dites de panel, mesures d'écarts entre échantillons de cas et échantillons témoins.

Plus de deux cents personnes ont assisté au colloque. Y sont intervenus les scientifiques et les experts auteurs de ces mesures, de grands-témoins acteurs du monde du travail – entreprises, syndicats – ainsi que les principales institutions compétentes – la Commission Nationale de l'Informatique et des Libertés, l'Agence nationale pour la Cohésion Sociale et l'Égalité des chances, le Centre d'Analyse Stratégique, la Mission d'animation de la recherche du Ministère du Travail, le Bureau International du Travail, la Haute Autorité de Lutte contre les Discriminations et pour l'Égalité.

Le présent rapport constitue les actes scientifiques du colloque. Il reprend les outils qui ont été distribués aux participants ce jour-là – le glossaire, les fiches signalétiques des enquêtes, les références disponibles en ligne –, et il reprend les transcriptions des principales présentations faites lors du colloque sur différents exercices statistiques et de testing.

Le colloque du 22 octobre 2007 à Lyon s'est tenu dans le cadre du projet AVERROES (Actions Visant l'Égalité sans distinction de Race de Religion ou d'Origine dans l'Emploi et les Services) qui s'est lui-même développé dans le cadre du programme européen Equal. Le colloque a par ailleurs reçu le label « 2007 - Année européenne de l'égalité des chances pour tous ». Il a été organisé par ISM CORUM.

Glossaire

Eric Cédiey, ISM CORUM, Benoît Riandey, INED

Un **recensement** est une opération statistique dans laquelle l'information est recueillie sur tous les individus d'une population ¹.

À l'opposé, dans une **enquête par sondage**, seule une partie de la population est enquêtée, mais sélectionnée avec une méthodologie (le sondage précisément) qui permettra l'extrapolation des résultats à l'ensemble de la population.

Si cette méthodologie a été mise en œuvre avec succès, nous sommes en présence d'un **échantillon représentatif** qui permet l'extrapolation (la généralisation) des résultats de l'enquête, par exemple l'estimation de la fréquence du phénomène étudié dans la population entière. [*cf. ci-après les enquêtes exposées par Julien Pouget et par Christelle Hamel*].

Par un choix historique de vocabulaire curieux et peu pédagogique, la méthode scientifique de sélection d'un échantillon représentatif est l'**échantillon aléatoire**. Les enquêtés y sont sélectionnés par tirage au sort, pas nécessairement avec les mêmes probabilités, mais des probabilités connues. Cette méthode est souvent longue, coûteuse et difficile à mettre en place.

Les **échantillons sur quotas** sont une autre méthode, empirique, pour constituer des échantillons visant à l'extrapolation des résultats à la population entière. Les enquêtés ne sont plus tirés au sort, ce qui génère un risque de biais des résultats. Le respect de certaines proportions (en général relativement à l'âge, au sexe, aux professions) sert de garde-fou à l'encontre du risque de biais. Ce sont précisément les quotas. L'expérience des instituts de sondage permet dans la plupart des cas de satisfaire au critère de représentativité.

Le fait de disposer par ailleurs de statistiques sur la population entière permet de rétablir dans l'échantillon les proportions correspondantes. C'est l'opération statistique, tout à fait légitime, du **redressement** de l'échantillon. Un redressement ne suffit pas toujours à rendre représentatif un échantillon biaisé.

Le **traitement d'un fichier de gestion** complet (fichier du personnel, administratif, etc.) présente les mêmes caractéristiques qu'un recensement. Néanmoins, on peut aussi opérer un sondage dans un fichier de gestion volumineux, en général pour suivre et traiter plus facilement l'échantillon extrait.

Une enquête par sondage au cours de laquelle on suivra un échantillon de population, réinterrogé plusieurs fois, s'appelle un **panel**. Le plus souvent, on complète régulièrement un panel statistique pour qu'à chaque collecte l'échantillon soit représentatif de la population du moment : on ajoute par exemple des immigrants ou des nouveaux-nés.

¹ A ce titre, le recensement rénové de la population réalisé par l'INSEE chaque mois de janvier n'est pas à proprement parler un recensement puisqu'on opère par sondage dans les communes de plus de 10.000 habitants. Cette remarque lexicale n'enlève rien à la pertinence statistique du recensement rénové.

Cette définition du mot panel n'a rien à voir avec l'évocation d'un groupe de personnes rassemblées pour donner leur point de vue sur un thème donné. La définition statistique du mot panel n'est pas strictement respectée non plus dans l'expression commune "comparaison de panel" (voir plus loin). Cette formule désigne une technique de mesure de la discrimination qui utilise la logique des enquêtes cas-témoins (voir plus loin) et non la logique des enquêtes par sondage.

Une **cohorte** est une population (définie par échantillonnage ou de manière exhaustive) constituée d'individus ayant en commun d'avoir connu un même événement durant une période donnée (entrée dans un pays, sortie du système éducatif) et qu'on réinterroge ensuite à des dates fixes (une forme de panel) ou sur lesquels on collecte des informations au fur et à mesure des événements qu'ils connaissent. [cf. *ci-après les enquêtes exposées par Thomas Coupié et par Marie-Noëlle Decharne et Eric Liedts*].

Les **échantillons raisonnés** constituent une autre logique ne visant pas à l'extrapolation à toute la population, mais plutôt à l'exemplarité. Une forme type en est la **monographie** : la sélection d'une ou plusieurs entreprises par exemple. Cette sélection sera opérée pour rendre compte de la diversité des situations, sans que la fréquence qui y sera observée soit généralisable à l'ensemble de la population (par exemple les salariés de toutes les entreprises en France). La monographie peut elle-même emprunter la forme du recensement, du sondage représentatif ou de l'échantillon raisonné. [cf. *ci-après l'enquête exposée par Patrick Simon*].

Un **plan d'expérience** est une forme d'échantillon raisonné, ne visant en général pas à l'extrapolation à la population complète, mais destiné à tester et démontrer l'effet d'un facteur donné. L'exemple type en est l'efficacité d'un médicament comparé à un placebo. Un tel essai pharmaceutique sera probant même en l'absence d'un échantillon représentatif de la population totale.

Le principe décisif est la comparabilité de l'échantillon des **cas** (recevant le médicament) avec celui des **témoins** (recevant à leur insu un placebo), les deux échantillons devant être de structure identique par ailleurs (d'où l'expression toutes choses égales par ailleurs). Ce type d'exercice sera donc appelé **enquête cas-témoins** et est destiné à prouver une causalité et non à mesurer une fréquence (prévalence d'une maladie au sein de la population). L'échantillon représentatif de la population ne sera pas en général l'échantillon le plus efficace pour prouver l'hypothèse de causalité testée.

La technique du testing (pour montrer l'effet d'un facteur ou critère donné sur les embauches par exemple), tout comme la technique dite de la "comparaison de panel" (par exemple pour montrer l'effet d'un critère sur les déroulements de carrière), relèvent toutes les deux de la logique de preuve de l'enquête cas-témoins.

Un **testing** consiste à ce que des personnes, qui sont en tous points comparables et qui ne diffèrent que sur un critère discriminatoire donné (leur sexe, leur origine supposée, etc.), effectuent des démarches identiques auprès des mêmes individus ou organisations dont il s'agit de tester les pratiques. Si les comportements recueillis varient en fonction du critère, l'influence de ce dernier est prouvée et des discriminations sont repérées.

Mais, vue la difficulté à constituer des échantillons de testés et de testeurs qui soient représentatifs d'une population réelle (par exemple celle des offreurs et des demandeurs d'emploi présents sur tel marché du travail à telle période), un testing ne peut prétendre

mesurer la fréquence réelle des discriminations au sein d'une population. Un testing s'appuie plutôt, suivant une logique d'**enquête cas-témoins**, sur des paires de testeurs qui sont stéréotypés au regard d'un facteur possible de discrimination (et non pas représentatifs de toute une population) ; le testing établit ainsi si des discriminations se produisent ou pas, et si oui sur quels critères, sous quelles formes, dans quels contextes.

Le terme **auto-testing**, apparu récemment dans le langage commun, désigne un testing auquel l'organisation qui est testée procède elle-même (par exemple une grande entreprise), ou plus sérieusement qu'elle commande elle-même à un tiers opérateur qui construit le protocole, met en œuvre les tests et reste le seul à décider et à savoir où et quand ces tests ont lieu. Dans ce dernier cas on parlera plutôt de **testing sollicité**, comme on le dit en matière d'audit financier ou de notation sociale. [*cf. ci-après l'enquête exposée par Fabrice Foroni*].

Testings et testings sollicités sont des tests de causalité applicables aux sélections à l'entrée, par exemple les embauches. Pour tester l'effet éventuel d'un facteur discriminatoire dans les trajectoires, comme les déroulements de carrière, on peut utiliser la technique dite de la "**comparaison de panel**".

Il faut comparer les parcours d'individus qui pourraient être touchés par la discrimination à tester (les **cas**) avec les parcours d'individus qui ne sont pas a priori exposés à cette discrimination, et qui sont comparables aux premiers quant à leur situation de départ : embauche dans les mêmes années, au même niveau de formation, sur le même coefficient ou échelon, dans la même filière (les **témoins**, imprécisément appelés "le panel"). Si les parcours connaissent des écarts au détriment des premiers individus, l'influence du facteur discriminatoire est montrée. On peut essayer de repérer à quels moments la discrimination se manifeste, sous quelles formes, etc.

Cadre problématique

Que mesure-t-on ? Pour quoi ? Comment ?

Eric Cédiey, ISM CORUM

Les débats sur la question dite des « statistiques ethniques » ont été nourris durant l'année écoulée [2007]. Cela va sans doute continuer : un amendement à la Loi informatique et libertés est en discussion au Parlement [octobre 2007], amendement qui va préciser la possibilité que soient conduites, après autorisation de la CNIL, des études sur, je cite, « la mesure de la diversité des origines des personnes, de la discrimination et de l'intégration ».

Il est très important que ces questions de mesure soient exposées et débattues. Au Parlement, et ailleurs.

Les débats jusqu'à présent se sont surtout focalisés sur la question des catégories : quelles catégories peut-on ou ne peut-on pas, doit-on ou ne doit-on pas, construire en la matière. Cette question des catégories est tout à fait cruciale. Et nous allons bien sûr l'aborder dans ce colloque.

Mais en amont de la question des catégories il faut répondre à d'autres questions tout aussi cruciales, dont la question des catégories pourrait dépendre, et qui pourtant ne sont pas souvent tirées au clair dans les débats.

À commencer par la question de l'objet, la question du « quoi » : qu'est-ce que l'on prétend, au juste, mesurer ? L'amendement déjà cité se réfère à « la mesure de la *diversité* des origines des personnes, de la *discrimination* et de l'*intégration* ». Ces trois objets, si l'on entreprend de les définir sérieusement, se révèlent très différents entre eux – en dépit du fait que bon nombre de discours les confondent allègrement. D'autres objets encore apparaissent dans les débats, objets qui sont à nouveau différents des premiers, comme les identités ou les appartenances. Or différents objets vont appeler différentes techniques de mesure, et différentes catégories pertinentes pour les mesurer.

Ces différents objets vont également être associés à différentes finalités. C'est la seconde question fondamentale, qui n'est sans doute jamais suffisamment tirée au clair dans les débats. C'est la question du « pour quoi », en deux mots, qui est très liée, bien sûr, à la question du « quoi » : qu'est-ce que l'on prétend déduire, ou faire, de ce que l'on prétend mesurer ? À ce stade prenons un exemple.

Comparons les deux premiers objets cités, à savoir la mesure de la « diversité » par rapport à la mesure de la « discrimination » (il s'agira ici de *mesure* de la diversité, pas de ce que l'on appelle le management de la diversité, qui est un autre sujet). Cet exemple entre diversité et discrimination est emblématique, tant sont nombreux ces dernières années les discours qui confondent les deux, en estimant que la mesure de la diversité vaut comme mesure des discriminations, et que l'augmentation de la diversité équivaut à la réduction des discriminations. C'est faux, tant du point de vue des objets que des finalités.

Sur la question de l'objet, pour les discriminations, nous disposons de définitions précises et opérationnelles, qui sont censées, de plus, faire référence commune, puisqu'elles sont données par le droit. Ces définitions indiquent clairement que les discriminations sont à chercher dans les actes et les pratiques, les traitements et les dispositions, les processus opératoires. Voilà les objets d'une mesure de la discrimination.

Certaines techniques vont mettre en œuvre cette mesure de façon directe, comme les testings, qui permettent de repérer directement les traitements discriminatoires dans les recrutements, par exemple. D'autres techniques vont le faire de manière indirecte, en mesurant certains écarts, toutes choses égales par ailleurs, entre les situations professionnelles par exemple, et en essayant de repérer les moments ou les dispositions pratiques qui construisent ces écarts.

Pour ce qui est de la finalité de ces mesures, le droit, à nouveau, pose la norme commune : les discriminations sont interdites. Si donc on tente de repérer les pratiques discriminatoires, c'est parce qu'il s'agit de les corriger et de travailler à leur disparition. Ce que l'on peut à nouveau tenter de mesurer.

Saisir les discriminations, puis les éliminer, n'est pas forcément facile. Mais l'objet et l'objectif sont connus.

En matière de diversité, il ne semble pas que nous bénéficions d'une définition qui fasse référence commune. Le plus souvent, on comprend néanmoins que la *mesure* de la diversité passerait par une identification et un classement des personnes en fonction de critères possibles de discrimination, par un décompte des personnes classées dans les différents groupes constitués selon ces critères, par une mesure de la distribution de la population sur ces différents groupes.

Les discours sur la mesure de la diversité – quand ils n'utilisent pas ce terme en lieu et place de discrimination – ne remontent guère aux pratiques. Les objets d'une mesure de la diversité, ce sont les personnes, les groupes, les populations.

En terme de finalité, l'objectif est alors très souvent formulé dans les termes de ce qu'on peut appeler la « théorie du reflet », selon laquelle les proportions des différents groupes de personnes à l'intérieur d'une organisation donnée, telle une entreprise ou une institution, devraient tendre à refléter les proportions externes dans la société.

Comme nous n'avons de définition et de norme commune que sur la discrimination, les discours sur la diversité présentent souvent ce reflet comme un gage de non-discrimination. Ce qui n'a aucun fondement, ni de droit, ni de fait. La reproduction, à l'interne, des proportions externes, ne garantit nullement qu'il y ait égalité de traitement dans les entrées et les sorties d'une organisation, ni à l'intérieur de celle-ci. Inversement, le non-reflet de ces proportions ne prouve aucunement qu'il y ait discrimination.

Nous manquons cruellement non seulement d'un cadre de référence, mais même de toute norme, aujourd'hui en France, pour pouvoir dire ce que serait un « bon » ou un « mauvais » niveau de diversité. Nous ne savons donc pas où nous mène ce discours.

Mais nous pouvons déjà constater qu'il peut mener à n'importe quoi. Je vais citer un cas, certes caricatural. Le problème, c'est que nous n'avons pas besoin d'inventer la caricature, elle est là. Et que ce cas ne fait que pousser au paradoxe la logique du reflet telle qu'on a pu l'entendre évoquée ces derniers temps en France.

Un homme politique français s'est autorisé à déclarer, parlant de l'équipe nationale de football, je cite : « Dans cette équipe, il y a neuf Blacks sur onze. La normalité (sic) serait qu'il y en ait trois ou quatre. Ce serait le reflet de la société ».

Cette assertion ne convaincra personne qu'il faut considérer que les systèmes de sélection dans le football professionnel, en France et ailleurs, qui mènent certains joueurs français jusqu'à leur équipe nationale, sont des systèmes violemment discriminatoires contre « les Blancs ».

Notons en même temps que la théorie du reflet, et la mesure de la diversité, ne nous offrent pas d'argument à rétorquer à notre homme politique. Elles sont au contraire constitutives du problème qu'il nous pose.

Et relevons au passage que ce n'est pas un débat sur les catégories admissibles qui peut nous offrir, non plus, quelque argument que ce soit : si, au lieu de dire « Blacks », notre homme politique dit « nés de parents eux-mêmes nés en Afrique sub-saharienne ou aux Antilles », le problème reste entier.

Ce type de problème indique qu'on ne peut faire l'économie de réfléchir et d'exposer très rigoureusement, dans nos débats et dans nos exercices de mesure, les objets que l'on prétend se donner et les finalités que l'on prétend poursuivre. Ce sont les réponses aux questions du « quoi » et du « pour quoi » qui donnent son sens, *ou pas*, à une mesure.

Vient alors la question du « comment », la question de la méthode. Celle-ci, dans ses divers éléments, doit être cohérente avec le sens donné à la mesure. Sur les thèmes qui nous occupent, trois éléments de méthode, au moins, vont varier selon les objets et les finalités :

- la logique statistique que réclame la mesure,
- les périmètres pertinents où l'appliquer,
- les catégories à construire.

Reprenons rapidement l'exemple de la comparaison entre diversité et discrimination.

Premier élément, la logique statistique. Pour bien mesurer la « diversité » d'une population, il faut compter ou représenter tous ses membres. Soit on recense et on classe tout le monde, soit on travaille sur un échantillon représentatif à partir duquel on pourra extrapoler les effectifs des différents groupes au sein de la population complète.

Pour repérer des pratiques ou des processus discriminatoires, il n'est pas besoin de classer tout le monde ou de se fonder sur des échantillons représentatifs de toute la population. On sera plus efficace en s'appuyant sur des échantillons stéréotypés au regard du critère dont on veut tester s'il donne lieu, ou pas, à des discriminations, dans telles pratiques ou dans tels processus. Le dénombrement de toutes les personnes susceptibles de pâtir de ces discriminations est une autre question, à laquelle il n'est pas nécessaire de répondre pour repérer les pratiques ou processus discriminatoires, et pour les corriger.

Benoît Riandey, dans la présentation qui va suivre, va exposer deux logiques différentes de mesure plus précisément que je ne viens de le faire.

Deuxième élément de méthode, le périmètre. L'échelle pertinente d'une mesure de la diversité, des « origines » par exemple, est celle de la population que l'on veut caractériser de la sorte. L'étendue de cette population peut être très variable. Il est possible, si l'on veut, d'entreprendre une mesure de la diversité des « origines » au sein de toute la population vivant en France à tel moment. Reste à savoir ce qu'il sera bon d'en penser et d'en faire.

Si l'on veut repérer des discriminations, alors il faut pouvoir rattacher l'exercice de mesure à un ensemble de pratiques identifiables : par exemple – dans le domaine de l'emploi – des procédures d'orientation, des procédures de recrutement, un système de formation, des processus d'insertion professionnelle, un dispositif d'accompagnement vers l'emploi, les démarches de gestion des ressources humaines d'un employeur donné, ou la formation des salaires... En revanche, ce que serait une mesure des discriminations dans toute « la société française » est on ne peut moins clair.

Troisième élément de méthode, la construction des catégories. La question apparaît délicate en ce qui concerne ce qu'on a nommé « l'origine », de façon très générale, dans le titre de ce colloque. Il faut là aussi être cohérent par rapport à l'objet et à la finalité de la mesure, et par rapport aux autres éléments de méthode.

Dans les phénomènes de discriminations, ce qui apparaît déterminant, ce sont les supports des perceptions d'autrui et les stéréotypes que ces perceptions mobilisent. Le législateur français en a parfaitement pris la mesure lorsqu'il a introduit le patronyme et l'apparence physique dans la liste des critères discriminatoires. Il ne semble pas que les discriminations liées à « l'origine » se fondent bien souvent sur des enquêtes quant à l'état civil des parents ou sur les sentiments des victimes elles-mêmes. Tout indique qu'elles s'en tiennent communément à des stéréotypes sur les consonances ou sur les apparences, sans s'enquérir des ascendances « réelles » ou des appartenances « authentiques ». Il est donc cohérent de procéder de même pour construire les catégories qui serviront à repérer ces discriminations.

Bien sûr, la prise en compte des prénoms et/ou des noms, *par exemple*, ne permet pas d'identifier de façon fiable « l'origine vraie » des personnes. Mais nous venons de voir que ce n'est pas le propos. Elle ne permet pas non plus de classer tout le monde. Mais nous avons vu que ce n'était pas nécessaire.

Deux logiques de mesure : enquêtes par sondage et représentativité, ou enquêtes cas-témoins et test de causalité

Benoît Riandey, INED

Je propose de distinguer et d'examiner deux logiques de mesure :

- la mesure de la fréquence de la discrimination,
- la preuve quantitative de l'existence d'une discrimination ou test probatoire de la discrimination.

Ces deux logiques de mesure, de la fréquence et de l'existence, vont chacune requérir un outil statistique différent :

- pour la mesure de la fréquence de la discrimination, ce sera l'enquête représentative de la population ;
- pour prouver l'existence de la discrimination, on sera plutôt dans le cadre du test d'expérience, de l'expérimentation, du testing.

La mesure de la fréquence de la discrimination

Quand on cherche à mesurer la discrimination, la première question qui va se poser sera : est-on en train de mettre en évidence des inégalités ou des discriminations ?

Prenons le cas des différences de salaires entre les hommes et les femmes. Dans les années soixante, où la qualification, le niveau d'instruction des femmes sur le marché du travail étaient très inférieurs à ceux des hommes, où les femmes les plus instruites étaient beaucoup plus jeunes, en moyenne, que les hommes de même niveau d'instruction, les différences de salaires pouvaient apparaître, avant tout, comme des différences structurelles et l'évidence d'une discrimination n'était pas si nette que cela. Aujourd'hui où, dans tous les niveaux de qualification, les femmes sont au moins aussi instruites que les hommes, les différences de salaires prouvent beaucoup plus fortement la discrimination.

Donc quand on cherchera à montrer la fréquence d'une discrimination, on aura à faire ce que les épidémiologistes appellent le « contrôle de tous les facteurs de confusion », c'est à dire de toutes les différences de structures que j'ai évoquées tout à l'heure, comme les différences d'âge et de niveau d'instruction, tous ces points qui peuvent expliquer les différences. Les statisticiens seront amenés à chercher les écarts qui sont constatés après avoir éliminé les effets de toutes ces causes structurelles, autres que le facteur discriminatoire lui-même. Il s'agit de faire une mesure « toutes choses égales par ailleurs ».

Est-ce facile ? Ça ne l'est pas tant que ça parce que quand on aura présenté son résultat « toutes choses égales par ailleurs », il y aura toujours un aimable collègue pour dire : « Mais vous avez oublié tel facteur, et c'est lui qui explique la différence que vous constatez ». Il y aura toujours ce que j'appelle la hantise de la variable cachée.

Ces efforts de méthode étant faits, il va falloir opérer la mesure, et pour cela, la première condition, ce sera obligatoirement qu'il va falloir observer, relever le caractère qui est la cause de la discrimination, que ce soit un patronyme étranger, un prénom étranger, la couleur de la peau qui n'est pas majoritaire dans le pays.

Dans les années passées, on a rencontré beaucoup d'obstacles pour pouvoir relever ces informations qui étaient indispensables à la mesure de la discrimination. Pendant toutes mes années passées au service des enquêtes de l'INED, j'ai eu bien souvent l'occasion de voir qu'on nous demandait de mesurer certains phénomènes sensibles, et que, en même temps, il y avait des freins très importants qui nous empêchaient de collecter les informations qui permettaient de mesurer ce qu'on nous demandait. Cela fait partie d'un paradoxe que beaucoup de statisticiens ont rencontré.

Les témoignages de la discrimination comme indicateur de mesure ?

Par rapport à cette obligation de mesurer le phénomène discriminatoire, une objection nous est présentée. Est-il vraiment nécessaire de collecter l'information auprès des personnes discriminées ? Le nombre de témoignages que l'on peut recueillir sur les discriminations n'est-il pas tout simplement une mesure de la discrimination ? L'objection paraît tout à fait recevable dans un premier temps, mais en fait elle est fragile. En effet, quand s'accroît la fréquence des témoignages, est-ce que ce sont les discriminations qui s'accroissent, ou la sensibilité aux discriminations ? La relation entre ces trois phénomènes n'est pas simple.

De plus, si l'on veut entreprendre une enquête statistique pour mesurer les discriminations à partir de témoignages, il faut une clef de passage de la fréquence des témoignages à la fréquence des discriminations réelles. Il faut au moins connaître le nombre de témoins de ces discriminations.

Or, les témoignages recueillis dans une enquête ne vont pas permettre de mesurer la fréquence des discriminations réelles, parce que les témoins en présence ne sont pas ceux qui permettent de faire cette mesure. Pour un conflit familial, père, mère, sœurs, frères, enfants, sont dénombrés, et ils peuvent donc être des témoins statistiques puisqu'on va pouvoir pondérer leur témoignage par l'inverse du nombre de témoins. Par contre, les amis, les voisins, les passants, les collègues, les militants, eux, ne sont pas dénombrés, donc leurs témoignages ne pourront jamais être pondérés par l'inverse du nombre de témoins de la discrimination dont on veut mesurer la fréquence. On ne peut donc pas passer de la fréquence des témoignages à la fréquence des discriminations.

La mesure de la fréquence des discriminations ne pourra donc être fondée que sur un échantillon représentatif de la population toute entière, parmi lequel on distinguera les personnes dotées du caractère susceptible d'induire la discrimination.

Prouver l'existence des discriminations

J'en viens maintenant à l'autre logique, celle du test probatoire : prouver qu'il y a une discrimination. Je prends un exemple tout à fait simple dans le domaine de l'épidémiologie : est-ce que l'aspirine soigne la migraine ? La méthode s'appelle le « plan d'expérience » : on va constituer deux groupes exactement semblables de migraineux ; dans l'un, on va donner de l'aspirine, dans l'autre, un placebo, bien sûr sans dire ce que l'on donne pour éviter un effet psychosomatique, et on compare la fréquence des satisfaits, des gens qui considèrent que ça va mieux. C'est la même chose que l'on fait dans les opérations de testing.

Y a-t-il une discrimination à l'embauche selon les origines ? On tente de constituer deux échantillons, qui ne vont différer que par un caractère, le choix du patronyme, le choix du prénom, ou le choix de la couleur de la peau, par exemple ; on constitue ces deux échantillons à partir de paires de quasi sosies qui ne diffèrent entre eux que par le caractère susceptible d'entraîner de la discrimination. Et on va mesurer, par exemple, le taux de retour des courriers, en comparant le groupe des « cas » (l'échantillon des individus porteurs du caractère discriminable) avec le groupe des « témoins » (leurs quasi sosies qui ne portent pas ce caractère).

La statistique nous fournit l'effectif nécessaire au test. Si l'écart est significatif, la discrimination est prouvée, sinon il n'y a pas de discrimination dite significative.

Quelques recommandations sur cette méthode :

1. Il faut avoir le test le plus puissant possible, et donc on va prendre des cas bien typés, bien typiques. Par exemple, il ne s'agirait pas, sur les prénoms, d'opposer, dans les « cas », Tania, et dans les « témoins » possibles, Nadia, la différence entre les origines des prénoms ne serait pas si typique que la tendance discriminatoire, si elle existe, puisse se manifester au maximum. Au contraire, il faudra comparer par exemple entre Fatia et Martine.

2. Nous n'avons pas besoin d'avoir un échantillon représentatif, en la circonstance, des prénoms. Nous cherchons au contraire à avoir des prénoms qui s'opposent bien, du point de vue des origines qu'ils évoquent, et non pas une distribution des prénoms représentatifs de ceux de l'ensemble de la population.

3. La fréquence de la discrimination qui va être observée par ce test ne sera donc pas celle qui existe dans l'ensemble de la population. Précisément, si vous prenez les prénoms les plus typiques, la discrimination va pouvoir s'opérer pleinement et donc plus souvent sans doute que sur les prénoms qui apparaissent moins typés. Donc dans un test efficace, la différence observée sera sans doute plus élevée que celle qu'il y aurait dans la population toute entière. On pourra conclure sur le fait qu'il y a une discrimination dans les situations auxquelles auront été appliqués les tests, mais on ne pourra pas extrapoler sur la fréquence globale de la discrimination.

Pour conclure, nous avons donc deux types d'opérations envisageables, la mesure de la fréquence des discriminations et le test de discrimination, qui sont des opérations statistiques légitimes et complémentaires l'une de l'autre.

Cadre réglementaire

La CNIL et la mesure des discriminations liées à « l'origine »

Sophie Vulliet-Tavernier, CNIL
directrice des affaires juridiques et de l'expertise

Quand on parle d'études sur la mesure de la diversité ou sur les discriminations, dans bien des cas, il y a traitement de données à caractère personnel, et la loi informatique et libertés s'applique. Je voudrais rapidement évoquer le panorama des règles applicables de protection des données, tout particulièrement en matière de données dites sensibles, puis les dix recommandations de la CNIL sur « la mesure de la diversité » adoptées en mai 2007.

La loi informatique et libertés et les données à caractère personnel

La loi informatique et libertés du 6 janvier 1978 a été modifiée assez profondément le 6 août 2004. Ces modifications se sont traduites par un renforcement des pouvoirs de contrôle et de sanction de la CNIL, en même temps que par un allègement des formalités déclaratives. Dans certains cas, et notamment dans le traitement de données sensibles, elles ont conduit à des régimes de contrôle a priori beaucoup plus lourds.

Par données à caractère personnel, on entend toute donnée qui permet d'identifier directement ou indirectement une personne physique. Dans bon nombre de traitements statistiques, dans bon nombre d'études, les traitements restent indirectement nominatifs, et sont soumis aux dispositions de la loi informatique et libertés, même si le nom des personnes n'est pas traité. Initiales du nom et du prénom, date de naissance ou autres indications de dates, adresse postale peuvent permettre d'identifier les personnes, et la loi informatique et libertés s'applique alors. Pour la CNIL, l'anonymat est réellement assuré quand il n'y a vraiment rien, même par corrélation de plusieurs données, qui puisse permettre d'identifier les personnes physiques. Cela sous-entend aussi de s'appuyer sur un échantillon de population suffisamment large, ou d'avoir recours à des techniques d'anonymisation telle que l'utilisation d'algorithmes de transcodage de l'identité – ces techniques développées dans le secteur de la santé mériteraient de l'être dans le cadre des études en sciences sociales.

La notion de traitement renvoie à toute opération sur des données, partant de la phase de collecte – notamment collecte par voie de questionnaire, y compris sur support papier – jusqu'à la diffusion des résultats, en passant par l'enregistrement, l'organisation, la conservation, la modification, l'extraction, la consultation, l'utilisation, la communication, le rapprochement, l'interconnexion, le verrouillage, l'effacement, la destruction : la loi informatique et libertés a donc une portée extrêmement large.

Les cinq règles d'or de la protection des données

1. Définition et respect de la finalité. Eric Cédiey a évoqué précédemment la notion de finalité. Lorsque la CNIL examine les projets, notamment dans le domaine de la recherche, elle formule une appréciation au cas par cas, sur la légitimité des objectifs poursuivis, en tenant notamment compte de la description précise des finalités de la recherche telles que décrites notamment dans le protocole.
2. Pertinence des données traitées. Ensuite nous apprécions, toujours au cas par cas, la nature et la pertinence des données collectées en regard de la finalité du traitement.
3. Conservation limitée des données. La durée de conservation des informations doit être limitée pour préserver le droit à l'oubli.
4. Sécurité et confidentialité. L'aspect sécurité et confidentialité est particulièrement important dans le domaine de la recherche et des statistiques, et appelle le développement de procédures de cryptage des informations, d'anonymisation, d'accès sécurisé aux données.
5. Respect des droits des personnes. En l'espèce, la loi impose un devoir de transparence aux organismes de recherche qui recueillent et traitent les informations. Il y a une obligation d'information des personnes enquêtées, qui porte non seulement sur l'exposé de la finalité du traitement, le caractère facultatif ou obligatoire des réponses, le droit de s'opposer à l'utilisation de leurs données, le droit d'accès et de rectification.

Le principe d'interdiction de traiter les données sensibles

L'article 8 de la loi informatique et libertés transposant la directive européenne d'octobre 1995, a introduit un régime très contraignant en ce qui concerne le traitement des données sensibles, puisque le principe introduit est celui de l'interdiction, avec toutefois un certain nombre d'exceptions. Les données sensibles sont celles qui font apparaître, directement ou indirectement, les origines raciales ou ethniques des personnes, les opinions politiques, philosophiques ou religieuses, l'appartenance syndicale, ainsi que les données relatives à la santé et à la vie sexuelle. Leur traitement est interdit, exception faite d'un certain nombre de dérogations.

S'agissant des questions liées aux « origines raciales ou ethniques », nous considérons que la nationalité, l'adresse et le lieu de naissance ne sont pas des données sensibles au regard de l'article 8. Ces données font tout de même l'objet d'une attention particulière de la CNIL, notamment dans le cadre de l'examen des projets de recherche qui nous sont soumis. Mais par exemple, nous avons parfaitement admis que ces informations figurent dans le tronc commun des enquêtes de l'INSEE.

En revanche, la CNIL considère comme relevant de l'article 8 :

- Les données sur les caractéristiques physiques des personnes, telles que la couleur de la peau. Dans certaines enquêtes sur le ressenti des discriminations, il peut y avoir des questions sur la couleur de la peau, et dès lors qu'il y a collecte de données permettant d'identifier les personnes, l'article 8 s'appliquera. Certains fichiers de

recherche médicale, notamment en dermatologie, peuvent également collecter des informations sur la couleur de peau. Certains fichiers de recrutement et certains fichiers de police – STIC, RG – ont été autorisés à enregistrer ce type d'information. Pour le recensement dans les TOM, il peut y avoir des informations sur l'appartenance tribale.

- Le traitement des patronymes peut faire apparaître l'appartenance des origines supposées. En l'espèce, la CNIL a rappelé que, dès lors qu'il est procédé à l'analyse de patronymes, cela relevait des conditions dérogatoires introduites par l'article 8. Nous avons d'ailleurs refusé en 2006 une enquête téléphonique qui devait être réalisée par le CRIF, et qui devait se baser sur l'analyse de la consonance des patronymes, supposée révéler une appartenance à la communauté juive. Nous n'avons pas estimé que cette méthode d'enquête était, en l'espèce, pertinente.

Les dérogations au principe d'interdiction

Quelles sont aujourd'hui, en l'état actuel de notre droit, les dérogations applicables au traitement de données sensibles dans le domaine de la recherche et des statistiques?

Consentement exprès des personnes concernées. En recherche et statistique, le chercheur peut obtenir le consentement exprès, c'est à dire écrit, des personnes concernées. Dans ce cas, il peut être autorisé à recueillir des données sur leur origine raciale. La contrepartie, si je puis dire, c'est que le fichier de recherche est soumis à un simple régime de déclaration auprès de la CNIL. Cette situation n'est à notre sens pas satisfaisante parce que, du point de vue de la méthodologie d'enquête, ceci oblige les enquêteurs à faire signer un formulaire de consentement aux enquêtés, ce qui est assez contradictoire avec le principe de l'anonymat et celui d'une relation de confiance qui doit s'établir entre l'enquêteur et l'enquêté. Par ailleurs, dans des enquêtes se déroulant dans le milieu de travail, on ne peut pas considérer que le consentement est réellement libre et éclairé. Ceci nous semble donc être une fausse garantie.

Anonymisation des données à bref délai et autorisation de la CNIL. Deuxième dérogation possible, si les données font l'objet d'une anonymisation à bref délai, et s'il y a autorisation de la CNIL. Mais sur ce point, dans bon nombre de cas, notamment pour les études sur la mesure de la diversité, il est nécessaire de conserver les noms ou des données indirectement nominatives, donc ce type de dérogation ne peut pas être appliqué.

Intérêt public et autorisation ou avis de la CNIL. Troisième cas de dérogation possible, l'enquête présente un intérêt public. Cette notion est toutefois délicate à apprécier. S'agissant de l'enquête de l'INED sur l'analyse de l'intégration des deuxièmes générations d'origine turque et marocaine, la CNIL a estimé, après discussion, compte tenu notamment de l'environnement européen de ce projet, et du fait que l'enquête était portée par l'INED, qu'elle présentait un intérêt public. Pour autant, qualifier systématiquement toute enquête sur la mesure de la diversité d'intérêt public, semble difficile.

INSEE ou services statistiques ministériels et autorisation de la CNIL. Dernière dérogation que j'évoquerai ici, les traitements statistiques réalisés par l'INSEE et les services statistiques ministériels, après avis du Conseil National de l'Information Statistique, sont également soumis à l'autorisation de la CNIL dès lors qu'ils portent sur des données sensibles.

Les dix recommandations de la CNIL en matière de « mesure de la diversité »

Je voudrais maintenant rappeler brièvement les dix recommandations de la CNIL adoptées en mai 2007. Nous avons estimé nécessaire de les rendre publiques parce que nous avons déjà adopté, en 2005, un premier rapport sur « la mesure de la diversité des origines », qui semble-t-il a été mal compris. Certains malentendus devaient être dissipés, la CNIL étant parfois perçue comme un frein au développement de la recherche et des études dans ce domaine.

Pour ce faire, il nous a paru nécessaire d'approfondir notre réflexion en la matière, de mieux comprendre les enjeux et les débats en cours. Nous avons donc engagé une série d'auditions avec l'ensemble des acteurs concernés, les chercheurs, les statisticiens, les syndicats, les associations, les églises et les mouvements religieux, etc. Soixante-dix auditions ont conduit à l'adoption d'un rapport, diffusé sur notre site, accompagné d'une série de recommandations, de propositions d'évolutions de notre loi et des pratiques sur les questions de mesure de la diversité. Ces recommandations n'ont pas de valeur normative, ce sont des lignes directrices avancées par la CNIL en la matière. Elles sont également disponibles sur notre site.

1. Ouvrir plus largement aux chercheurs l'accès aux bases statistiques existantes et aux fichiers de gestion, notamment ceux de l'Education Nationale, des organismes de protection sociale, évidemment dans le respect de la loi. On constate encore trop souvent que des bases de données statistiques, je pense à celles de l'INSEE issues du recensement, sont trop fermées. Nous reprenons les suggestions faites par le directeur général de l'INSEE lui-même et par Roxane Silberman, secrétaire générale du Comité de concertation pour les données en sciences sociales, qui prônent le développement, comme dans d'autres pays, de centres d'accès sécurisés, qui permettraient à des chercheurs d'avoir accès à des données plus individuelles, dans le respect de la confidentialité, et à des données plus fines issues du recensement. Nous avons engagé une concertation avec l'INSEE pour revoir les conditions de diffusion de ces données, s'agissant notamment des données qui, jusqu'à présent, étaient sujettes à des restrictions particulières de diffusion, comme les données sur la nationalité, le pays de naissance, l'année d'arrivée en France, ou la notion d'immigré. La CNIL est ouverte à une redéfinition des modalités de diffusion de ces données.

2. Utiliser, dans les enquêtes, les données relatives à l'ascendance des personnes, non seulement dans le cadre d'enquêtes spécifiques (et la CNIL l'a déjà admis dans le cadre de quelques enquêtes), mais pourquoi pas aussi dans les questionnaires du recensement. Nous sommes favorables à une étude de faisabilité de l'intégration de ce type de question dans le recensement.

3. Mais ne pas intégrer de données sur l'ascendance dans les fichiers des administrations et des entreprises. Il y a en effet des risques de mauvaise compréhension, de rupture de la confiance des administrés ou des salariés qui pourraient suspecter une utilisation détournée de ces informations.

4. Développer des études sur le ressenti des discriminations, pouvant inclure des questions sur l'apparence physique des personnes, mais plutôt dans le cadre de la statistique publique. Avec une validation scientifique préalable des questionnaires par une instance scientifique, et après des tests d'acceptabilité de ces questions, en informant bien sûr les enquêtés du caractère facultatif des réponses. Et en obtenant pour finir l'autorisation de la CNIL, ce qui peut ne pas être le cas actuellement.

5. Admettre, sous certaines conditions, l'analyse des noms et/ou des prénoms pour détecter des discriminations. La CNIL y est favorable pour détecter des discriminations, mais opposée dès lors qu'il s'agirait de classer ces noms et prénoms dans des catégories ethno-raciales. La CNIL recommande une catégorisation des noms et prénoms sous les modalités « discriminant / non discriminant », et elle recommande leur utilisation dans le cadre d'analyses multicritères prenant en compte la profession, le sexe, l'âge, etc.

6. Modification de la loi informatique et libertés pour prévoir un régime d'autorisation de la CNIL pour toute étude sur la mesure de la diversité. Lorsque la CNIL a adopté cette recommandation, il s'agissait de prévoir un régime d'autorisation uniforme pour tous les traitements de données sensibles, dès lors qu'ils sont mis en œuvre à des fins de recherche et de statistique. Actuellement, il existe déjà, dans la loi, un régime d'autorisation spécifique pour les fichiers de recherche médicale, après avis d'un comité scientifique. Ce que nous souhaiterions, c'est élargir à l'ensemble des traitements de données sensibles ce régime d'autorisation systématique par la CNIL, après, si nécessaire, une expertise scientifique des projets déposés. Cette expertise pourrait s'appuyer sur le Comité de concertation pour les données en sciences sociales, créé en octobre 2001.

7. Pas de référentiel ethno racial. La CNIL était déjà clairement réservée sur ce point en 2005, et les auditions que nous avons menées ont montré que de façon générale, il y avait un consensus sur l'idée qu'il devait y avoir une démarche progressive et prudente en la matière. De toute façon, ouvrir la possibilité de constituer un tel référentiel incomberait au législateur, sous le contrôle du Conseil Constitutionnel. La position de la CNIL reste extrêmement réservée sur ce point.

Les trois dernières recommandations portent sur des points plus techniques :

8. Recourir à des experts tiers de confiance, type sociétés d'études, etc., qui, pour la CNIL, sont les plus à même de mener ce genre de mesures en garantissant la confidentialité des données et l'anonymat des résultats, y compris vis-à-vis des commanditaires des mesures en question.

9. Recourir à des techniques d'anonymisation, j'ai déjà évoqué ce point précédemment.

10. Assurer la transparence et mieux garantir les droits des personnes, parce que je pense que dans le monde de la recherche, il y a vraiment des progrès à faire sur l'information des personnes concernées par les études.

Enquêtes par sondage en population générale

Enquête FQP et enquête Structure des Salaires – Études sur la diversité dans la fonction publique et sur les différences salariales dans le secteur privé

Julien Pouget, INSEE

Merci de m'avoir invité pour présenter ces travaux, menés conjointement depuis trois ans avec une petite équipe de recherche (Romain Aeberhardt, Denis Fougère, Roland Rathelot), au CREST. En écho à l'introduction d'Eric Cédiey à ce colloque, je ne sais pas si ces travaux traitent plutôt de diversité ou de discriminations, et je crois que je vais suivre un continuum entre les deux : en partant d'un constat statistique qui est celui de la diversité, ou de son éventuelle absence, pour ensuite essayer de montrer dans quelle mesure des méthodes permettent, ou non, de mettre en évidence des phénomènes de discrimination. Ces travaux portent sur l'accès à l'emploi et les différences salariales, et s'appuient sur des enquêtes menées à l'INSEE tant auprès des ménages que des entreprises.

Je commencerai par rappeler succinctement les problématiques que nous avons été amenés à traiter, avant de décrire les sources, en rappelant, comme l'a fait Benoît Riandey, les avantages et les limites de ces enquêtes en populations générales ou sur échantillons représentatifs. Je décrirai brièvement les deux sources que nous avons utilisées, c'est à dire l'enquête sur la Formation et la Qualification Professionnelle de 2003, menée auprès des ménages, et l'enquête sur la Structure des Salaires de 2002, menée principalement auprès des entreprises, avec un sous-échantillon constitué en parallèle auprès des salariés. J'en viendrai brièvement aux questions de méthode, puis donnerai un aperçu des résultats de nos travaux sur ces enquêtes.

1. Problématique

Etudier l'insertion sur le marché du travail nécessite de disposer d'un certain nombre de données, d'informations, sur le capital humain initial des personnes, et notamment leur niveau d'éducation, dont on sait qu'il conditionne, pour une grande partie, l'accès à l'emploi. On peut alors essayer dans un premier temps d'observer s'il existe ou non, dans l'accès à l'emploi, des phénomènes discriminatoires. Puis, à l'emploi est attaché un revenu d'activité, et l'on va également essayer d'étudier s'il existe ou non des discriminations touchant les carrières ou les salaires.

Que nous permettent de faire les différentes sources de données ? L'enquête sur la Formation et la Qualification Professionnelle (FQP) nous donne un panorama de l'ensemble de la population, elle concerne à la fois les actifs, les chômeurs, les salariés, les indépendants, mais aussi les inactifs. On va donc pouvoir, à partir de cette source, disposer d'informations sur les déterminants de l'accès à l'emploi. On a également utilisé cette enquête pour étudier plus spécifiquement les déterminants de l'accès à l'emploi public, et tenter d'expliquer les différences salariales.

Par ailleurs, on dispose de l'enquête sur la Structure des Salaires menée auprès des entreprises sur un échantillon de salariés uniquement, de taille plus conséquente, et avec des informations plus précises sur leurs salaires. Cette enquête contient, comme l'enquête FQP, des variables socio-démographiques de capital humain, mais on ne va pas pouvoir ici mesurer les déterminants de l'accès à l'emploi, puisque par définition les enquêtés sont déjà salariés. En revanche, on va pouvoir tenter de déterminer les facteurs qui peuvent favoriser l'accès au statut de cadre, et ceux qui peuvent expliquer les écarts salariaux.

Avec de telles données, on n'observe pas le détail des mécanismes, des processus qui ont pu mener les personnes à un emploi ou au chômage (encore que, dans l'enquête FQP, on dispose d'un petit historique). On n'observe que les *résultats* de ces processus, avec néanmoins l'avantage de disposer de données représentatives, qui concernent l'ensemble des secteurs. La limite, c'est que les différences que je vais qualifier d'« inexplicables » peuvent être dues à de la discrimination mais également à d'autres facteurs. Autrement dit, quand on veut contrôler un ensemble d'effets de structure, on est tributaire, précisément, des variables de structure dont on dispose, qui peuvent influencer la partie inexplicée ou « résiduelle » des différences, partie qui peut éventuellement relever de phénomènes discriminatoires. Et par ailleurs ces enquêtes ne présentent pas d'informations fines sur les mécanismes de discrimination, ce qui peut réduire la portée du discours quant aux leviers d'action publique à mettre en œuvre par la suite.

En terme de catégories statistiques, on se situe dans le cadre que rappelait la CNIL tout à l'heure, c'est-à-dire des enquêtes de l'INSEE, qui partent de la définition du terme « immigré » que je rappelle ici pour mémoire : une personne née étrangère, à l'étranger, ce qui implique donc des conditions sur le lieu de naissance et sur la nationalité à la naissance. Les variables qui sont collectées dans les enquêtes évoquées plus haut sont la nationalité et le lieu de naissance non seulement des enquêtés, mais aussi de leurs parents. Dans les travaux que je suis amené à résumer ici, on va distinguer les personnes selon les lieux de naissance et nationalités à la naissance de leurs parents. On sera souvent amené à distinguer trois sous-populations : les personnes n'ayant aucun parent immigré, celles qui en ont un et celles qui en ont deux. Beaucoup d'études montrent en effet que les personnes qui ont un parent immigré ont des caractéristiques et des probabilités d'accès à l'emploi et des salaires relativement proches, en définitive, des personnes qui n'en ont aucun. La différence la plus pertinente à retenir dans l'analyse est sans doute celle entre les personnes n'ayant aucun ou un parent immigré, et les personnes ayant deux parents immigrés.

2. Sources

Les sources de données utilisées sont décrites dans les fiches qui ont été distribuées aux participants du colloque [« Fiches signalétiques » reprises pages 30 et 31 à la suite du présent exposé]. Il s'agit d'enquêtes qui sont toutes deux les éditions les plus récentes de séries déjà relativement anciennes, puisque les premières éditions datent des années soixante, mais qui comprennent pour la première fois des questions sur le lieu de naissance et/ou la nationalité à la naissance des parents.

L'enquête FQP de 2003 est échantillonnée sur l'ensemble des personnes âgées de 18 à 65 ans, elle est réalisée en face à face, avec un échantillon final d'environ quarante mille répondants. Elle comprend donc les variables sur le lieu de naissance et la nationalité à la naissance des

parents. A partir de cette enquête, on a exploré les déterminants de l'accès à la fonction publique, et plus généralement l'insertion sur le marché du travail et les différences salariales.

L'enquête sur la Structure des Salaires est une enquête européenne, c'est-à-dire avec un tronc commun de variables communes collectées dans l'ensemble des pays européens. Néanmoins la variable qui nous intéresse ici plus particulièrement, le pays de naissance des parents, n'a pas été collectée dans l'ensemble des pays de l'Union : il s'agit d'une spécificité de l'enquête conduite en France.

C'est une enquête relativement originale puisqu'elle a un double, voire un triple questionnaire. On envoie un questionnaire postal auprès des entreprises pour collecter un ensemble d'informations, non seulement sur les entreprises, leur secteur, etc., mais également sur le salaire d'un sous-échantillon de leurs salariés. On va avoir ici des données sur les salaires qui seront beaucoup plus précises que ce qu'on peut obtenir en mode déclaratif dans le cadre d'une enquête auprès des ménages. En parallèle, on envoie au domicile de ce même sous-échantillon de salariés un questionnaire léger, comprenant un certain nombre de questions sur leur situation socio-démographique, dont le pays de naissance de leurs parents. Et l'INSEE réalise l'appariement entre les deux questionnaires. A partir de cette enquête sur la Structure des Salaires, on ne va pas pouvoir étudier des thématiques comme l'accès à l'emploi. On va plutôt se focaliser sur les différences salariales, les données collectées étant très précises en la matière, et sur l'accès au statut de cadre.

3. Méthodes

Brièvement et sans trop entrer dans la technique, que peut-on utiliser comme méthodes, en allant du plus simple au plus sophistiqué, et, peut-être, de la diversité à la discrimination ?

Dans un premier temps, on peut former des moyennes, ou des médianes, et les comparer entre différents groupes.

Ensuite, on va essayer de faire des choses plus fines en contrôlant les effets de structure : les taux d'emploi observés dans les différentes sous-populations ne sont pas les mêmes, leurs niveaux d'éducation ne sont pas les mêmes non plus, et on va pouvoir expliquer un certain nombre de choses grâce notamment à des variables comme l'expérience professionnelle ou le niveau d'éducation. On va alors proposer des décompositions plus fines, qui partagent les différences salariales, par exemple, entre une composante expliquée par ces effets de structure, par ces caractéristiques observables de niveau d'éducation ou d'expérience, et une composante non expliquée, qui va mesurer en quelque sorte les différences de rendement des caractéristiques précédentes. Par exemple, une année d'éducation en plus produit 5 à 6% de surcroît salarial. Est-ce que ce rendement de 5 à 6% est le même au sein des différentes sous-populations ? C'est ce que l'on va être amené à étudier.

Et puis si on veut aller encore un petit peu plus loin, on va réaliser cette même décomposition des différences salariales, mais en prenant également en compte les effets de sélection, c'est-à-dire le fait que les gens ne se répartissent pas aléatoirement entre les différents groupes (cadres, non cadres, salariés, chômeurs). Or les sélections dans l'accès à l'emploi, ou dans l'accès au statut de cadre, ont bien sûr des effets propres sur les différences salariales.

4. Résultats

Voici, brièvement, quelques résultats, qui portent au départ simplement sur des comparaisons de moyennes, donc des résultats statistiques assez bruts. Dans le tableau suivant, on constate que le taux de chômage des personnes n'ayant aucun parent immigré était dans l'enquête FQP de 2003 de 9,4%, à comparer à 17,9% pour les personnes ayant deux parents immigrés, et 30,1% pour les personnes ayant deux parents immigrés originaires du Maghreb. Ce dernier chiffre est entre parenthèses et en italique dans le tableau, parce qu'il reste fragile : on se heurte en effet rapidement à la faiblesse des échantillons de population - c'est une des limites de cette enquête - et donc les chiffres qu'on est amené à produire dans ce cadre sont un petit peu plus fragiles que ce qu'on aurait obtenu sur des populations plus importantes.

Part des actifs, taux de chômage, part des salariés du public et du privé en fonction du nombre de parents immigrés

En %

	Part des actifs dans la population		Taux de chômage	Part des salariés du public parmi les actifs occupés
	Totale	Hors étudiants et retraités		
Aucun parent immigré	73,7	89,0	9,4	24,5
Un parent immigré	71,2	86,3	11,9	22,5
dont :				
<i>un parent immigré originaire d'un pays d'Europe</i>	72,1	85,6	11,0	23,5
<i>un parent immigré originaire du Maghreb</i>	(69,5)	(88,5)	(15,2)	(18,9)
Deux parents immigrés	69,6	87,4	17,9	16,5
dont :				
<i>deux parents immigrés originaires d'un pays d'Europe</i>	74,0	88,0	11,4	16,3
<i>deux parents immigrés originaires du Maghreb</i>	(64,5)	(85,4)	(30,1)	(17,1)

Note : les chiffres indiqués entre parenthèses sont fragiles.

Champ : Français âgés de 18 à 65 ans, nés en France.

Source : Insee, enquête Formation et qualification professionnelle 2003.

Si on veut faire la part des effets de structure dans la probabilité d'avoir un emploi, on va être amené à expliquer le fait d'être en emploi ou au chômage par un certain nombre de caractéristiques, notamment le niveau d'éducation, la profession des parents, etc. On met alors en évidence, dans le tableau suivant, tant pour les hommes que pour les femmes, un effet marginal négatif du fait d'avoir des parents immigrés, plus précisément une diminution de la probabilité d'être en emploi de 2 à 3% pour les hommes, et de 4% pour les femmes. Mais ceci seulement pour les femmes qui ont deux parents immigrés, car en revanche on n'observe aucun élément significatif pour les femmes qui ont un seul parent immigré, c'est-à-dire que leur probabilité d'accéder à l'emploi, toutes choses égales par ailleurs, est à peu près équivalente à celles qui n'ont aucun parent immigré.

Impact des différentes caractéristiques des actifs sur leur probabilité d'avoir un emploi

		Hommes		Femmes	
		Coefficient	Effet marginal	Coefficient	Effet marginal
	Constante <i>Probabilité de la situation de référence</i>	1,47	81,2 %	0,89	70,8 %
Origine des parents	<i>Aucun parent immigré</i>	<i>référence</i>		<i>référence</i>	
	Un parent immigré	- 0,32	- 2,0	n.s.	
	Deux parents immigrés	- 0,43	- 2,7	- 0,51	- 4,4
Âge	<i>18-29 ans</i>	<i>référence</i>		<i>référence</i>	
	30-39 ans	0,85	+ 5,4	0,80	+ 6,9
	40-49 ans	1,02	+ 6,5	1,13	+ 9,8
	50-65 ans	0,37	+ 2,4	0,59	+ 5,1
Diplôme le plus élevé obtenu en formation initiale	<i>CEP, BEPC ou aucun diplôme</i>	<i>référence</i>		<i>référence</i>	
	BEP-CAP	0,60	+ 3,8	0,51	+ 4,4
	Baccalauréat	0,65	+ 4,1	0,80	+ 6,9
	Bac+2	1,04	+ 6,6	1,58	+ 13,7
	Supérieur	0,93	+ 5,9	1,29	+ 11,1
Profession du père	Inconnue ou sans objet	- 0,42	- 2,7	- 0,28	- 2,5
	Agriculteur	1,36	+ 8,6	0,99	+ 8,6
	Artisan, commerçant, chef d'entreprise	0,31	+ 2,0	n.s.	
	Cadre du public	0,55	+ 3,5	0,12	+ 1,0
	Cadre du privé	n.s.		- 0,35	- 3,1
	Profession intermédiaire du public	n.s.		0,50	+ 4,3
	Profession intermédiaire du privé	n.s.		n.s.	
	Employé ou ouvrier du public	n.s.		n.s.	
	Employé du privé	n.s.		n.s.	
	<i>Ouvrier du privé</i>	<i>référence</i>		<i>référence</i>	

n.s. : non significatif au seuil de 10 %.

Si on se focalise maintenant sur la probabilité non plus d'avoir un emploi mais, une fois qu'on est salarié, d'être dans la fonction publique, on met également en évidence, dans le tableau suivant, un effet marginal négatif du fait d'avoir deux parents immigrés, à peu près équivalent pour les hommes et les femmes. Là encore, on n'observe pas de différence significative entre les personnes qui n'ont qu'un seul parent immigré et celles qui n'en ont aucun.

Impact des différentes caractéristiques des salariés sur leur probabilité de travailler dans la Fonction publique

		Hommes		Femmes	
		Coefficient	Effet marginal	Coefficient	Effet marginal
	Constante <i>Probabilité de la situation de référence</i>	2,14	10,5 %	- 1,72	15,2 %
Origine des parents	<i>Aucun parent immigré</i>	<i>référence</i>		<i>référence</i>	
	Un parent immigré	n.s.		n.s.	
	Deux parents immigrés	- 0,40	- 6,5	- 0,31	- 6,6
Âge	<i>18-29 ans</i>	<i>référence</i>		<i>référence</i>	
	30-39 ans	0,31	+ 5,0	0,35	+ 7,6
	40-49 ans	0,75	+ 12,2	0,70	+ 15,1
	50-65 ans	0,86	+ 14,0	0,96	+ 20,8
Diplôme le plus élevé obtenu en formation initiale	<i>CEP, BEPC ou aucun diplôme</i>	<i>référence</i>		<i>référence</i>	
	BEP-CAP	n.s.		0,22	+ 4,7
	Baccalauréat	0,59	+ 9,6	0,53	+ 11,4
	Bac+2	0,34	+ 5,4	0,85	+ 18,5
	Supérieur	0,93	+ 15,0	1,46	+ 31,6
Profession du père	Inconnue ou sans objet	n.s.		n.s.	
	Agriculteur	n.s.		n.s.	
	Artisan, commerçant, chef d'entreprise	n.s.		- 0,32	- 6,9
	Cadre du public	0,48	+ 7,8	0,28	+ 6,1
	Cadre du privé	n.s.		- 0,52	- 11,1
	Profession intermédiaire du public	0,85	+ 13,7	0,29	+ 6,3
	Profession intermédiaire du privé	n.s.		- 0,15	- 3,2
	Employé ou ouvrier du public	0,74	+ 12,0	0,41	+ 8,9
	Employé du privé	0,25	+ 4,0	n.s.	
	<i>Ouvrier du privé</i>	<i>référence</i>		<i>référence</i>	

n.s. : non significatif au seuil de 10 %.

On peut également utiliser l'enquête FQP pour procéder aux types de décompositions que j'ai brièvement présentées plus tôt, en prenant en compte les effets de sélection. Le résultat principal, pour aller à l'essentiel, c'est que les variables de contrôle habituelles (âge, genre, niveau d'éducation, expérience, lieu de résidence, etc.) n'expliquent que les deux tiers des différences salariales, et seulement la moitié des différences dans l'accès à l'emploi. Autrement dit, on s'aperçoit que s'il y a des barrières, *elles se manifestent sans doute plus fortement en termes d'accès à l'emploi qu'en termes de différences salariales.*

On a été amené à mettre en œuvre une méthodologie voisine à partir de l'enquête sur la Structure des Salaires. Il s'agit non plus de modéliser la probabilité d'avoir un emploi, mais celle d'accéder au statut de cadre, et bien sûr d'évaluer les déterminants des salaires. Ceci sur trois sous-populations de salariés français à temps complet, dont les parents sont nés soit en France, soit en Europe du Sud, soit au Maghreb. On obtient deux résultats principaux.

Premier résultat : il apparaît que les rendements salariaux de l'expérience professionnelle globale d'une part, et de l'ancienneté spécifique dans l'entreprise d'autre part, diffèrent dans les différentes sous-populations étudiées. Plus précisément, on observe des rendements de l'expérience professionnelle globale plus importants pour les personnes dont les deux parents sont nés en France, mais en revanche des rendements de l'ancienneté spécifique dans l'entreprise plus importants pour les personnes dont les deux parents sont nés à l'étranger. Une interprétation possible peut être liée à ce que l'on nomme en économie la *discrimination statistique*. Les candidats dont les deux parents sont nés en Europe du Sud ou au Maghreb connaissent un moindre rendement de leur expérience professionnelle parce qu'à l'embauche ils sont en butte à une plus grande méfiance des employeurs ; autrement dit la barrière à franchir quand les employeurs ne connaissent pas encore les salariés potentiels qui proposent leurs services est plus élevée pour ces derniers que pour les candidats dont les deux parents sont nés en France. En revanche, une fois que les employeurs connaissent ceux qu'ils ont embauchés et à mesure que l'ancienneté de ces derniers s'accroît dans l'entreprise, ils valorisent d'autant plus les caractéristiques véritablement observées de leurs salariés qu'ils les avaient potentiellement sous-estimées au moment de l'embauche.

Le second résultat est mis en évidence dans le tableau suivant. Sur la première ligne du tableau, figurent les salaires des hommes dont les deux parents sont nés au Maghreb, comparés à ceux dont les deux parents sont nés en France : « -3,2 » signifie que l'écart brut, la différence des moyennes salariales, est de -3,2% en défaveur des salariés dont les parents sont nés au Maghreb. Une fois pris en compte les effets de structure et de sélection, on s'aperçoit que l'écart résiduel est en revanche positif, c'est à dire qu'on explique totalement la différence salariale initiale et qu'on va même au-delà. Autrement dit, il ne semble plus y avoir aucune discrimination salariale quand on prend en compte l'ensemble des effets de structure et de sélection.

Écart de salaire et de probabilité d'accès au statut de cadre

		Écart brut	Écart résiduel
Écart entre les hommes dont les deux parents sont nés au Maghreb et ceux dont les deux parents sont nés en France	Ensemble des salariés	- 3,2	3,0
	Cadres	8,2	11,8
	Non-cadres	- 1,2	2,1
	Part de cadres	- 4,6	- 1,1
Écart entre les hommes dont les deux parents sont nés en Europe du Sud et ceux dont les deux parents sont nés en France	Ensemble des salariés	- 6,8	2,1
	Cadres	- 5,7	- 0,1
	Non-cadres	1,8	4,0
	Part de cadres	- 10,0	- 2,0
Écart entre les femmes dont les deux parents sont nés au Maghreb et celles dont les deux parents sont nés en France	Ensemble des salariés	0,3	2,0
	Cadres	17,7	12,4
	Non-cadres	1,4	4,1
	Part de cadres	- 4,2	- 3,9
Écart entre les femmes dont les deux parents sont nés en Europe du Sud et celles dont les deux parents sont nés en France	Ensemble des salariés	- 0,7	3,1
	Cadres	- 5,7	- 2,7
	Non-cadres	1,2	2,3
	Part de cadres	- 1,3	2,2

C'est là le second résultat important de ce travail : la majeure partie des différences salariales provient d'effets de structure comme les différences de niveaux d'éducation, et d'effets de sélection comme les barrières à l'entrée du statut de cadre (voir les lignes « Part de cadres » dans le tableau précédent). Une fois que l'on a pris en compte tout cela il reste peu de choses, voire plus rien en termes de différences salariales pures. Les barrières se situent plutôt à l'entrée à l'emploi, à l'accès au statut de cadre.

Conclusion

Ces travaux nous ont donc permis de mesurer des écarts résiduels. Ces écarts résiduels peuvent éventuellement être dus à de la discrimination, mais nos résultats ne sont pas de la même nature que ceux qui peuvent résulter d'un *testing* par exemple. Un résultat semble important : les principales barrières se situent plutôt au niveau de l'accès à l'emploi et de l'accès aux professions les plus rémunératrices, mais une fois que ces barrières sont franchies, il semble n'y avoir que très peu de discrimination purement salariale. En termes de leviers de politiques publiques de lutte contre les discriminations salariales, on ne va donc pas être amené à mobiliser les mêmes outils pour réduire les écarts selon l'origine et les écarts selon le sexe. En effet, quand on développe la même modélisation pour étudier les différences salariales hommes-femmes, on s'aperçoit qu'il reste beaucoup de différences qu'on n'explique pas avec les variables dont on dispose ; alors que lorsque l'on réalise comme ici l'exercice sur les différentes sous-populations distinguées selon le lieu de naissance des parents, on explique quasiment la totalité des différences salariales. Dans ce dernier cas, les actions à mener doivent donc porter sans doute davantage sur l'éducation et sur l'emploi, avec une prise en compte des spécificités des recrutements selon les différents secteurs (l'exemple de la fonction publique, évoqué tout à l'heure, est sans doute parlant de ce point de vue).

Fiche signalétique

Intervenant : **Julien Pouget (Crest-Insee)**

Nom de l'opération : **Enquête sur la Structure des Salaires en 2002**

Source(s) des données (enquête spécifique, exploitation d'enquête(s) générale(s), exploitation de fichier(s) de gestion) : Enquête quadriennale recueillie par l'Insee, et coordonnée au niveau européen par Eurostat. Cette enquête comprend l'interrogation d'un échantillon d'établissements appartenant à des entreprises de 10 salariés ou plus, et l'interrogation directe d'un échantillon de leurs salariés.

Lieu(x) de la collecte : France métropolitaine.

Mode(s) de recueil (interne ou externe, auto-administré ou par enquêteur ; téléphone, face à face, internet, postal...) : Postal, auprès des établissements d'une part, et des salariés d'autre part.

Population ou périmètre concernés : salariés des entreprises de plus de 10 salariés, appartenant aux sections C à K de la nomenclature d'activités de la communauté européenne, c'est-à-dire l'industrie, la construction, le commerce et les services marchands à l'exclusion notamment de la santé.

Type et taille de l'échantillon, ou effectif de la collecte exhaustive : on dispose d'un échantillon de 120 000 salariés pour lesquels les entreprises ont transmis des informations. Mais le taux de réponse au questionnaire adressé directement aux salariés étant inférieur à celui du questionnaire adressé aux entreprises, les données socio-démographiques utilisées pour l'étude ne sont disponibles que pour environ 40 % de l'échantillon (soit un échantillon final d'environ 50 000 salariés).

Donnée(s) utilisée(s) pour construire les catégories d'"origines" : Pays de naissance des parents.

Principales questions explorées en lien avec les variables d'"origines" : différences salariales et différences dans l'accès au statut de cadre.

Principaux résultats : S'il existe des écarts importants dans l'accès à l'emploi et au statut de cadre, en revanche, « toutes choses égales par ailleurs », les salariés français dont les deux parents sont nés au Maghreb ou en Europe du Sud ne perçoivent pas des salaires significativement moins importants que leurs collègues dont les deux parents sont nés en France. Les résultats suggèrent ainsi que les éventuelles discriminations porteraient moins sur les salaires que sur l'accès à l'emploi ou aux catégories socioprofessionnelles les plus rémunératrices. Une fois ces barrières franchies, les salariés obtiendraient, à poste et niveau de qualification équivalents, des niveaux de salaires sensiblement égaux, quel que soit le pays de naissance de leurs parents.

Références :

Aeberhardt R. et J. Pouget (2006) : « Comment expliquer les disparités salariales », in Les salaires en France, édition 2006, collection Références, Insee, pp. 29-42.
http://www.insee.fr/fr/ffc/docs_ffc/SALFRA06ac.PDF

Aeberhardt R. and J. Pouget (2007): « National Origin Wage Differentials in France: Evidence from Matched Employer-Employee data », IZA Discussion Paper No 2779 and CREST Working Paper n°2007-3. <http://ftp.iza.org/dp2779.pdf>

Fiche signalétique

Intervenant : **Julien Pouget (Crest-Insee)**

Nom de l'opération : **Enquête sur la Formation et la Qualification Professionnelle (FQP) en 2003**

Source(s) des données (enquête spécifique, exploitation d'enquête(s) générale(s), exploitation de fichier(s) de gestion) : En 2003, l'Insee a réalisé la sixième enquête sur la formation et la qualification professionnelle (FQP). Initiée en 1964, l'enquête a aussi eu lieu, sous des formes proches, en 1970, 1977, 1985 et 1993. Ces enquêtes constituent une des principales sources d'information sur la mobilité professionnelle, la mobilité sociale et sur les relations entre la formation initiale et professionnelle, l'emploi et les salaires.

Lieu(x) de la collecte : France métropolitaine

Mode(s) de recueil (interne ou externe, auto-administré ou par enquêteur ; téléphone, face à face, internet, postal...) : en face à face.

Population ou périmètre concernés : personnes âgées de 18 à 65 ans.

Type et taille de l'échantillon, ou effectif de la collecte exhaustive : environ 40 000 répondants.

Donnée(s) utilisée(s) pour construire les catégories d'"origines" : lieu de naissance et nationalité de naissance des parents.

Principales questions explorées en lien avec les variables d'"origines" :

- Accès à la fonction publique.
- Différences salariales et différences dans l'accès à l'emploi.

Principaux résultats :

- Les personnes dont les deux parents sont immigrés, sur-représentées parmi les personnes sans emploi, sont moins souvent salariées de la fonction publique que celles qui n'ont aucun parent immigré. Cette sous-représentation semble légèrement plus importante que celle observée dans le secteur privé. Elle s'explique en partie par des différences de niveau d'éducation.
- Les variables de contrôle habituelles (âge, genre, niveau d'éducation, expérience, lieu de résidence, etc.) expliquent les deux tiers des différences salariales, et la moitié des différences dans l'accès à l'emploi, entre les salariés français dont les deux parents étaient français à la naissance, et ceux dont au moins un des deux parents avait, à la naissance, la nationalité d'un pays d'Afrique. Ceci suggère l'existence de discriminations, concernant notamment l'accès à l'emploi.

Références :

Pouget J. (2005) : « La Fonction publique : vers plus de diversité ? », in *France, Portrait Social*, édition 2005-2006, Insee, pp. 143-162. http://www.insee.fr/fr/ffc/docs_ffc/FPORSOC05c.PDF

Fougère D. et J. Pouget (2004) : « L'emploi public s'est-il diversifié ? Sexe, niveau d'étude, origine sociale et origine nationale des salariés de la fonction publique », Complément au Rapport sur *La diversité de la fonction publique* remis par Madame Dominique Versini au Ministre de la Fonction Publique et de la Réforme de l'Etat, La Documentation Française, Paris, 60 pages. <http://lesrapports.ladocumentationfrancaise.fr/BRP/054000049/0000.pdf>

Aeberhardt R., Fougère D., Pouget J. and R. Rathelot (2007): « Wages and Employment of French Workers with African Origin », IZA Discussion Paper No 2898. <http://ftp.iza.org/dp2898.pdf>

L'enquête européenne TIES : The Integration of European Second Generation – Méthode d'échantillonnage : patronymes et screening téléphonique

Christelle Hamel, INED

Je vais rapidement vous présenter le protocole d'échantillonnage de l'enquête TIES (The Integration of European Second Generation), enquête qui s'est appuyée sur une méthode de sélection patronymique, c'est-à-dire de reconnaissance des patronymes par leur affiliation à une langue.

Le protocole général de l'enquête

TIES (<http://www.tiesproject.eu/index.php>) est une enquête européenne réalisée dans huit pays. Elle a été initiée par nos collègues néerlandais de l'IMES, notamment Maurice Crul. Patrick Simon (INED) et moi-même avons porté le volet français de l'enquête. Elle porte sur les trajectoires des descendants directs de migrants originaires de Turquie, nés dans les pays de l'enquête. Dans chaque pays, un échantillon comparatif est constitué de personnes nées dans le pays et dont les deux parents sont également nés dans ce pays. Sont pris en compte les individus résidant dans deux villes au sein de chaque pays, pour le cas de la France, il s'agit de Paris et de Strasbourg.

L'objectif de cette recherche est d'étudier les parcours d'intégration des individus, l'intégration étant entendue comme un processus sur le long terme, que l'on peut étudier sur plusieurs générations. L'objectif de l'enquête est d'identifier quel est l'impact des contextes nationaux et locaux sur les parcours d'intégration des descendants de migrants turcs, en comparant les résultats obtenus dans les différents pays afin de voir comment les politiques d'intégration propres à chaque pays influent sur les parcours des individus.

Cet objectif implique de recueillir des données liées au pays de naissance des enquêtés et de leurs parents, ce qui est particulièrement compliqué dans le contexte français. Nous avons envisagé plusieurs solutions d'échantillonnage avant de porter notre choix sur la méthode patronymique.

Les difficultés d'échantillonnage

Les difficultés d'échantillonnage sont liées au fait que la population cible est extrêmement rare au sein de la population française : il s'agit de personnes dont l'un ou les deux parents sont nés en Turquie, ayant entre 18 et 35 ans, et réparties sur les deux agglomérations de Paris et Strasbourg.

Nous avons exploré plusieurs possibilités d'échantillonnage qui chacune se sont révélées inutilisables :

- La première solution envisagée consistait à organiser une enquête filtre ad hoc au recensement afin de repérer le pays de naissance d'ego et de ses parents. Le recensement étant en phase de rénovation en 2007, il ne fut pas possible de procéder ainsi.
- La deuxième solution envisagée était de rechercher dans les grandes enquêtes qui avaient recueilli les informations sur le pays de naissance des parents des enquêtés s'il était possible de constituer un échantillon de descendants d'immigrés turcs. Mais dans l'enquête *Histoire familiale* conduite en 1999 auprès de 380 000 personnes, les descendants de migrants turcs ayant entre 18 et 59 ans ne sont que 344 répartis sur toute la France. Or, nous devions constituer un échantillon de 500 personnes. L'enquête étant ancienne, elle ne permettait de toute façon pas de repérer les personnes ayant 18 à 35 ans aujourd'hui.
- La dernière solution que nous aurions pu suivre était de faire passer un petit questionnaire cherchant à identifier le pays de naissance d'ego et de ses parents, à l'ensemble des personnes abonnées à France Télécom et donc repérables sur l'annuaire téléphonique. Cependant, pour toucher un enquêté dans notre cible, il aurait fallu passer un nombre si important de coups de téléphone que le coût de l'enquête était prohibitif.

C'est à ce stade que nous avons décidé d'utiliser la méthode de reconnaissance linguistique des patronymes. Nous avons eu recours à la méthode de sélection patronymique afin de constituer, à partir de l'annuaire téléphonique, une base de sondage plus restreinte que l'annuaire téléphonique lui-même.

Notre méthode d'échantillonnage s'est déroulée en deux temps.

1. Tout d'abord, nous avons appliqué la méthode de sélection patronymique aux pages blanches de l'annuaire téléphonique de 2006. Nous avons ainsi constitué un fichier de 10.500 noms ayant un lien probable avec la langue turque.
2. Nous avons ensuite administré à chacun de ces numéros de téléphone un petit questionnaire dans lequel nous demandions au répondant de décrire tous les membres de son ménage, de façon à identifier les personnes nées en France dont l'un ou les deux parents étaient bien nés en Turquie.

La sélection patronymique préalable à la passation du questionnaire sur les membres du ménage a permis de considérablement réduire les coûts de l'enquête en augmentant la probabilité de trouver des personnes correspondant aux caractéristiques recherchées.

La CNIL a donné son accord pour cette enquête, à condition que les enquêtés soient d'abord avertis par courrier de la méthode de sélection de leur nom dans l'annuaire. Au téléphone, des explications complémentaires ont été données aux enquêtés.

La sélection des noms dans l'annuaire

La sélection des noms a été réalisée par un institut allemand (Humpert und Schneiderheinze) spécialisé dans l'étude généalogique des patronymes, leur transmission et leur diffusion. Cet institut travaille avec les services statistiques des universités et des instituts de recherche en

Allemagne. L'institut dispose de dictionnaires de noms et procède de deux façons pour collecter les noms qui peuvent être associés à une langue.

- Chaque langue a ses spécificités dans la manière de former ses noms propres. En Turquie, les noms les plus courants sont : des noms de métiers, des noms qui indiquent la filiation (c'est-à-dire que « fils de » ou « fille de » est inclus dans le patronyme lui-même), des titres honorifiques, ou bien réfèrent à des lieux ou à des références poétiques connues des spécialistes.
- Ensuite, pour les patronymes qui échappent à cette première catégorie facilement identifiable, on peut procéder à une analyse linguistique. L'institut repère les associations de consonnes et de voyelles, leurs enchaînements, les séquences phonétiques et établit une probabilité d'appartenance à la langue turque.

C'est ainsi qu'ils constituent leurs annuaires de patronymes. Dans leurs procédures, ils prennent également en compte les variantes existant lors de la transcription d'un nom étranger dans le pays où se réalise l'enquête.

Avec cette méthode, on ne procède pas à une imputation directe d'une origine migratoire à un nom, mais on établit une probabilité d'avoir une origine donnée.

Les biais de sélection liés à la méthode patronymique

Cette méthode comporte plusieurs biais de sélection qui la rendraient inutilisable pour sélectionner des individus dont les parents seraient originaires d'un autre pays que la Turquie. Ces biais sont au nombre de trois.

Premièrement, dans chaque pays, il existe des patronymes qui renvoient à un univers linguistique autre que celui du pays en question : ce sont les patronymes liés à des courants migratoires. Ainsi des personnes originaires de Grèce et installées depuis plusieurs générations en Turquie qui auraient migré vers la France pourraient se trouver exclues de notre base de sondage. Ce type de biais n'a en fait pas d'incidence en ce qui concerne la population originaire de Turquie. En effet, en 1934, sous le régime d'Ata Türk, une loi sur les patronymes a contraint les individus appartenant à des minorités linguistiques et migratoires à choisir un patronyme turc. Ce processus de répression des minorités (notamment juives, grecques et arméniennes) a obligé ces dernières à « turquiser » leurs patronymes. Dans le cas turc, la sélection patronymique par la méthode linguistique permettait donc de tenir compte des personnes issues de minorités au sein même de la Turquie.

Ensuite, les femmes immigrées qui se sont mariées avec un homme portant un nom français ne peuvent être repérées. L'incidence de ce biais est relativement faible appliquée à notre population, car il se trouve que les femmes immigrées turques sont assez peu nombreuses à avoir un conjoint né en France, c'est le cas de 4% d'entre elles seulement selon le recensement de 1999, et parmi ces 4%, le mari peut être descendant de migrants turcs, ce qui dans ce cas ne les exclut pas. Nous avons estimé, à partir de l'enquête *Histoire familiale*, que les personnes issues d'un couple mixte où c'est la femme qui est originaire de Turquie seraient 3,5% des descendants d'immigrés turcs. Le biais est donc relativement peu élevé, il aurait été beaucoup plus fort si on avait fait l'enquête sur les descendants de migrants marocains.

Le dernier biais est lié cette fois-ci au mariage des descendantes de migrants turcs. Si une jeune femme de parents immigrés de Turquie se marie avec une personne ne portant pas un nom turc, elle ne peut être identifiée. Mais là encore, l'incidence est relativement faible : toujours d'après l'enquête *Histoire familiale*, 80% des descendantes de migrants turcs se sont mariées avec un immigré turc, et 20% avec un conjoint né en France, mais ce dernier pouvant lui-même avoir des parents venus de Turquie, il n'est pas évident que parmi ces 20% toutes aient un nom qui ne soit pas turc.

Ainsi, les biais sont réels mais limités pour ce qui est de notre population.

Le résultat de cette sélection patronymique sur l'annuaire a permis de constituer un échantillon de 10 500 numéros de téléphones, sur un ensemble de 1 700 000 noms figurant sur l'annuaire pour les deux agglomérations. La proportion de ménages où personne n'est d'origine turque est d'environ 19%, confirmant l'efficacité de la méthode linguistique utilisée par cet institut.

La sélection téléphonique (screening téléphonique)

Une fois réalisée cette sélection des patronymes, les descendants de migrants n'ont pas été identifiés en tant que tels. A ce stade, nous avons donc procédé à un screening par appel téléphonique sur l'ensemble des noms de l'annuaire qui avaient été retenus. Ce screening téléphonique avait deux objectifs. Le premier était d'identifier la population cible en recueillant les informations sur le pays de naissance des individus et de leurs parents, leur âge et leur ville de résidence ; le deuxième objectif était aussi de corriger les effets des biais listés ci-dessus, même si ces derniers n'étaient pas trop importants.

Les biais que nous avons voulu corriger sont de deux sortes : l'absence des femmes mariées avec des personnes portant un nom qui n'est pas turc et l'absence des individus non inscrits dans l'annuaire téléphonique. Pour corriger ces biais, nous avons posé un grand nombre de questions sur l'ensemble des membres du ménage correspondant à chaque numéro de téléphone, sur leurs enfants, y compris ceux vivant en dehors du ménage.

Dans le cas où nous trouvions des jeunes femmes vivant dans ou en dehors du ménage, nous demandions quel était leur statut matrimonial. Si elles étaient mariées, nous demandions le pays de naissance du père de leur mari. Dans le cas où ce pays n'était pas la Turquie, la jeune femme portait donc probablement un patronyme autre que turc et elle devenait prioritaire pour l'enquête, ce qui permis de réintroduire des jeunes femmes ayant ce profil dans notre échantillon.

De la même façon, les personnes vivant en dehors du ménage et possédant exclusivement un téléphone portable devenaient prioritaires pour l'enquête (puisque'on ne pouvait les contacter par l'annuaire téléphonique).

Sur 10 550 numéros sélectionnés, environ 6600 ont abouti à un contact, les autres numéros n'étant plus bons ou correspondant à des entreprises ou bien étant toujours libres ou au contraire toujours occupés. Près de 2500 personnes ont accepté de répondre au screening, les autres déclarant d'emblée soit ne pas vouloir participer à l'enquête, soit que personne ne correspondait au profil recherché dans leur ménage. De nombreux ménages présentaient au moins une personne éligible pour l'enquête, mais un seul individu a été sélectionné par

ménage. Nous avons ainsi sélectionné 863 personnes. Certaines, dont les coordonnées avaient été transmises par le répondant au screening, sont restées injoignables, d'autres ont refusé l'enquête. Nous avons cessé les appels des personnes éligibles lorsque nous avons atteint les 500 entretiens en face-à-face.

Conclusion

Si la méthode patronymique ne permet pas à elle seule d'identifier les descendants d'immigrés turcs, sa combinaison avec la procédure de screening téléphonique s'est révélée efficace. Mais il faut souligner que cette efficacité tient aux caractéristiques de la population cible. Les biais inhérents à cette méthode nous ont ainsi empêchés de faire l'enquête auprès des descendants d'immigrés marocains, car d'une part les patronymes marocains sont indissociables des patronymes algériens, et d'autre part les femmes immigrées marocaines sont très fréquemment mariées avec des hommes non-marocains, si bien que leurs enfants nés en France n'auraient pu être enquêtés, alors qu'ils constituent une part importante des descendants d'immigrés marocains. Cette méthode est donc un outil utile mais elle ne peut être utilisée que de façon circonstanciée.

Fiche signalétique

Intervenante : **Christelle Hamel (Ined)**

Nom de l'opération : **Enquête TIES : The Integration of European Second generation**

Source des données : Enquête spécifique

Lieu(x) de la collecte : Cette enquête internationale est réalisée dans 16 villes de huit pays d'Europe : l'Allemagne, l'Autriche, la Belgique, l'Espagne, la France, les Pays-Bas, la Suisse, la Suède. En France, l'enquête est réalisée dans les agglomérations de Paris et Strasbourg

Mode de recueil : Enquête par questionnaires en face à face, menée par un institut de sondage.

Population ou périmètre concernés : En France, l'enquête est conduite auprès des descendants directs de migrants originaires de Turquie, et de personnes nées en France descendantes directes de personnes également nées en France. Tous les pays du projet TIES conduisent l'enquête auprès de descendants de migrants turcs, auprès d'un échantillon témoin de natifs du pays dont les parents sont également nés dans ce pays, et auprès d'un second groupe minoritaire : les descendants de migrants marocains ou les descendants de migrants ex-yougoslaves selon l'histoire migratoire du pays. En France, des raisons budgétaires ont empêché de conduire l'enquête auprès des descendants de migrants originaires du Maroc.

Type et taille de l'échantillon, ou effectif de la collecte exhaustive :

En France, dans chacune des deux agglomérations, sont interrogées :

- 250 personnes, âgées de 18 à 35 ans, nées en France dont l'un ou les deux parents sont nés en Turquie
- 200 personnes, âgées de 18 à 35 ans, nées en France dont les deux parents sont nés en France

Donnée(s) utilisée(s) pour construire les catégories d'"origines" :

La procédure d'échantillonnage de l'enquête TIES s'appuie sur la reconnaissance patronymique des origines nationales et a reçu un avis favorable de la CNIL en mai 2007. Le choix de cette méthode s'est imposé par l'impossibilité de repérer les descendants de migrants originaires de Turquie par d'autres moyens. En effet, le recensement étant en phase de rénovation, il n'était pas possible de procéder à une enquête filtre associée au recensement pour identifier la population cible ; et les grandes enquêtes identifiant le pays de naissance des parents de l'enquêté, qui auraient pu servir de base de sondage, sont trop anciennes (EHF date de 1999) et ne contiennent pas d'effectifs suffisants dans les villes où l'étude est conduite. Appliquer le questionnaire de screening aux ménages ayant un numéro sur l'annuaire téléphonique, sans sélectionner au préalable ces ménages par le biais de la méthode patronymique, aurait nécessité de passer un nombre d'appels si élevé que cette hypothèse n'était matériellement pas réalisable.

La procédure d'échantillonnage des descendants de migrants originaires de Turquie se décline donc en deux étapes :

1/ Repérage sur l'annuaire téléphonique dans les agglomérations de Strasbourg et Paris des patronymes ayant une consonance les rapportant à la langue turque. Cette sélection a été opérée par la société allemande Humpert und Schneiderheinze, spécialisée dans les recherches onomastiques, et appliquée sur l'annuaire de 2006. Cette sélection a permis de repérer 7056 noms ayant une forte probabilité d'être portés par des individus nés en Turquie ou ayant un père ou un grand père originaire de ce pays. La Turquie ayant mis en œuvre une politique de « turquisation » des patronymes des groupes minoritaires résidant sur son territoire sous le régime d'Atatürk, ce mode de sélection par le patronyme n'induit pas l'occultation des minorités comme les kurdes par exemple. Appliquée à une population originaire d'autres pays où la diversité linguistique des patronymes est plus forte, la méthode patronymique peut s'avérer tout à fait inopérante. En l'occurrence, elle a permis de constituer la base de sondage pour la constitution de l'échantillon.

2/ Un screening téléphonique a ensuite été appliqué à ces 7056 numéros de téléphone. Ce screening consiste en la passation d'un questionnaire court décrivant l'ensemble des individus résidant dans le logement du numéro d'appel. Il permet d'identifier le pays de naissance de chaque individu et de chacun de ses parents, ainsi que son âge, son sexe, son niveau de diplôme. Ces renseignements ont permis de sélectionner la population ciblée par l'enquête : les 18-35 ans nés en France dont l'un des parents au moins est né en Turquie. L'échantillon ainsi constitué est représentatif de la population des descendants de migrants originaires de Turquie disposant d'un numéro de téléphone fixe. L'analyse des biais et des possibilités de correction de cette méthode d'échantillonnage sont en cours.

La procédure d'échantillonnage des personnes nées en France de parents également nées en France consiste en la passation du questionnaire de screening sur une liste de numéros de téléphone choisis aléatoirement sur l'annuaire 2006 dans les deux villes étudiées.

Principales questions explorées en lien avec les variables d'"origines" :

Expérience des discriminations dans la trajectoire scolaire et l'emploi.

Construction sociale de l'identité et liens avec le pays d'origine.

Principaux résultats :

Les premiers résultats en France feront l'objet d'un rapport qui devrait être finalisé pour la fin de l'année 2008. L'enquête TIES a un site Internet dédié : « TIES: L'Intégration de la Deuxième Génération en Europe » : <http://www.tiesproject.eu/index.php?lang=fr> (description du projet d'étude et de son réseau européen, programme des recherches, publications, etc.)

Analyses de cohortes

Différenciations des parcours professionnels des jeunes selon l'origine socioculturelle

Thomas Couppié, CEREQ

Je vais présenter de quelle façon les travaux du CEREQ permettent d'approcher les questions de l'origine et des discriminations dans le champ de l'insertion professionnelle des jeunes. Les travaux du CEREQ autour de l'insertion s'appuient en particulier sur les enquêtes « Génération », qui sont des enquêtes biographiques sur des cohortes sortant du système éducatif.

Trois points seront abordés :

1. Les enquêtes « Génération »
2. La façon dont les origines sont abordées dans ces enquêtes
3. La question des discriminations liées à l'origine, et leur mesure dans ces enquêtes.

1. Les enquêtes « Génération »

Elles ont pour vocation de caractériser les relations formation-emploi en début de vie active, en retraçant les parcours professionnels des jeunes sortant des établissements de formation initiale. Quatre générations de sortants du système éducatif ont pour l'instant été enquêtées, les générations de 1992, de 1998 que nous analyserons ici, et les générations de 2001 et 2004.

Les enquêtes « Génération » s'attachent à recueillir des données sur le type d'études achevées et sur la situation professionnelle des jeunes durant des périodes d'observation variables : trois, cinq ou sept années de vie active. La collecte peut se faire au cours d'une interrogation unique, ou de plusieurs interrogations successives : c'est le cas de la génération 1998 qui a été interrogée en 2001 puis en 2003 et 2005. Les données collectées sont essentiellement factuelles, décrivant les différentes situations rencontrées par les jeunes, et un calendrier professionnel récapitulant leur parcours est constitué. Mais on collecte également des questions d'appréciation, de jugement de la part des jeunes : on demande par exemple aux jeunes si, dans une situation donnée, ils s'estiment bien payés, s'ils s'estiment employés à leur niveau de compétence, etc.

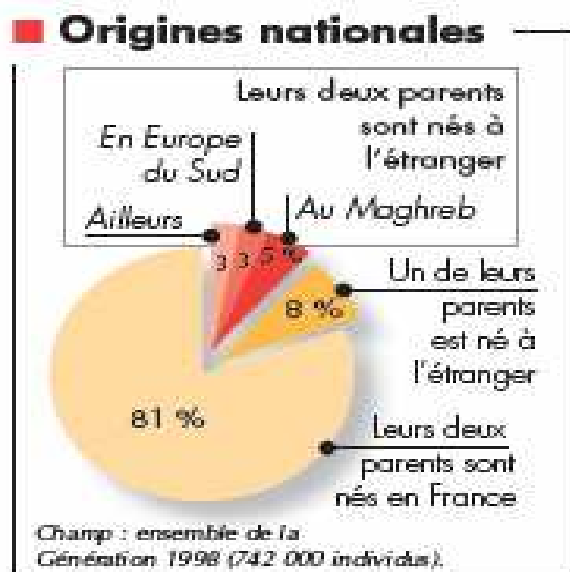
2. La place des origines dans ces enquêtes

Pourquoi vouloir mesurer les origines dans un tel dispositif ?

Premièrement, parce que comme d'autres critères de stratification sociale, tels que le genre ou la catégorie sociale, l'origine est un critère qui est susceptible d'interagir avec le diplôme quand on s'intéresse à l'insertion des jeunes. Deuxièmement, parce que, comme d'autres critères de stratification sociale, le critère de l'origine joue doublement dans la détermination des positions atteintes, en amont et en aval de l'entrée dans la vie active, c'est à dire sur :

- les positions atteintes dans le système éducatif, qui conditionnent le niveau de diplôme auquel les jeunes vont accéder,
- leurs situations sur le marché du travail.

Le choix du type de définition de l'origine a d'abord répondu aux contraintes posées par la CNIL au milieu des années 90. Par ailleurs, le CEREQ souhaitait s'appuyer sur des nomenclatures simples, supposées fiables et pérennes, ne demandant pas d'investissement méthodologique. Enfin, nous voulions traiter la question de l'origine en tant que critère de stratification sociale et non comme une variable cible comme dans les enquêtes de mesure de la diversité par exemple. Nous avons donc opté pour la méthode généalogique : on saisit le lieu de naissance du jeune, le lieu de naissance de ses parents, et la nationalité de ses parents.



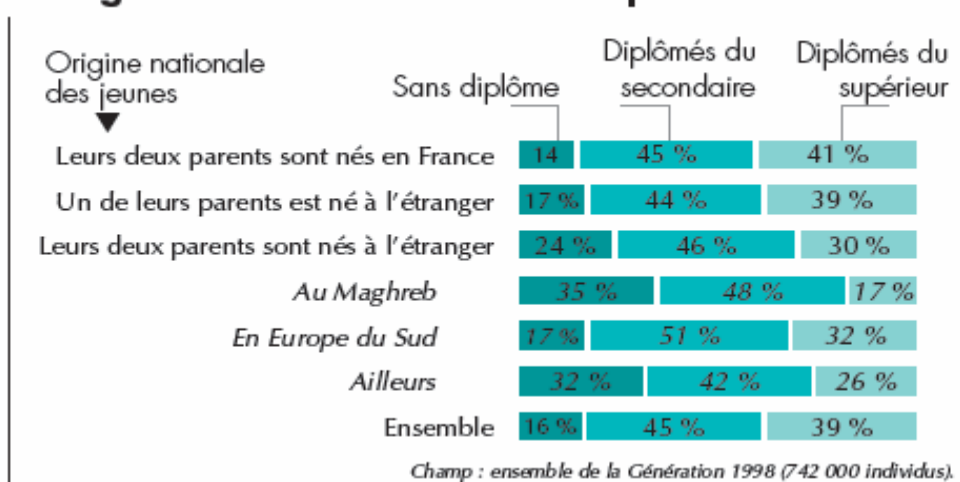
Source : Couppié, Gasquet, Lopez (coord.) : Quand la carrière commence..., Céreq, 2007.

Ces choix apparemment simples ne résolvent pas tout. Une fois ces informations rassemblées, différentes façon de les agréger sont possibles : par exemple, selon qu'on considère la naissance à l'étranger des deux parents ou d'un seul, on isole des populations qui peuvent aller de 11 à 19% de la génération. De même, un certain nombre de groupes peuvent être statistiquement distingués et notamment, parmi les jeunes dont les deux parents sont nés à l'étranger, ceux qui sont originaires du Maghreb et ceux originaires d'Europe du sud. Ces populations vont pouvoir être comparées avec le groupe des jeunes dont un seul des parents est né à l'étranger, et a fortiori avec le « groupe de référence » constitué des jeunes dont les deux parents sont nés en France.

Les principaux constats relatifs au poids de l'origine dans les parcours des jeunes

On constate d'abord que l'origine joue en amont de l'entrée dans la vie active, dans la détermination des positions atteintes dans le système éducatif. Par exemple, la probabilité de sortir diplômé de l'enseignement supérieur varie considérablement selon les origines : 41% des jeunes dont les deux parents sont nés en France sont diplômés de l'enseignement supérieur pour seulement 17% des jeunes dont les deux parents sont nés au Maghreb. Inversement, la proportion de non diplômés s'élève à 35% pour ces derniers, soit 2,5 fois plus que ce qu'on observe chez les jeunes dont les deux parents sont nés en France.

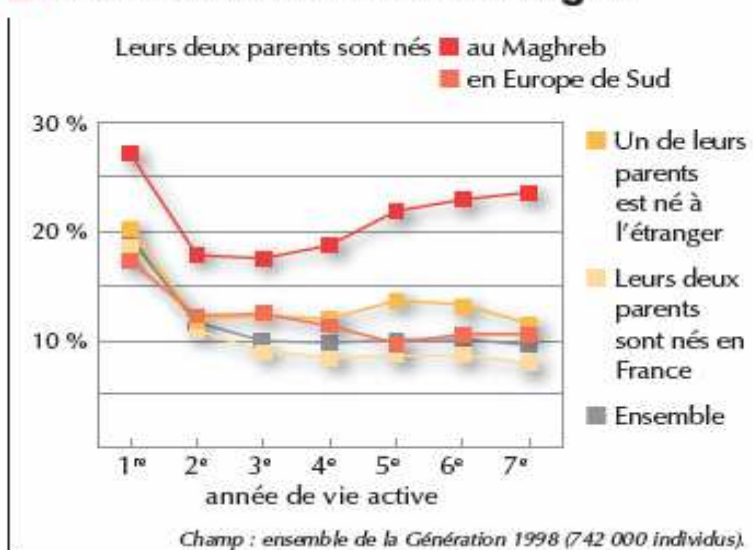
■ Origine nationale et niveau de diplôme



Source : Couppié, Gasquet, Lopez (coord.) : Quand la carrière commence..., Céreq, 2007.

On constate ensuite que ces origines ont un effet en aval quand on observe l'évolution des taux de chômage au fil des sept premières années de vie active. Les jeunes d'origine maghrébine interrogés dans l'enquête apparaissent particulièrement exposés au chômage, et ont particulièrement souffert du retournement de conjoncture de 2001.

■ Évolution du taux de chômage



Source : Couppié, Gasquet, Lopez (coord.) : Quand la carrière commence..., Céreq, 2007.

Une question se pose alors : peut-on inférer de la discrimination à partir de ces observations ?

3. La mesure des discriminations

Si on mesure bien à partir des enquêtes « Génération » des inégalités sociales, c'est à dire des écarts de situations entre des groupes différents, on ne peut pas prétendre, dans ce type d'enquête générale, mesurer des discriminations au sens juridique du terme. La raison en est assez simple : à aucun moment nous n'avons les moyens d'observer des interactions entre le

trio discriminateurs / discriminés / non discriminés – comme les procédures de testing le permettent, par exemple – parce que, par construction du dispositif, notre regard n'est que partiel et n'utilise qu'un seul point d'entrée, qui sont les individus sur le marché du travail.

Cela signifie-t-il que l'on ne peut rien dire ici au chapitre des discriminations ? Non, on peut mettre en évidence ce qu'on peut appeler des présomptions de discrimination. Deux méthodes peuvent être sollicitées dans le cadre des enquêtes « Génération ».

La modélisation économétrique

La méthode de la modélisation économétrique repose sur l'approche initiée par Oaxaca et Blinder il y a une trentaine d'années. Elle vise à distinguer, dans les inégalités observées (accès à l'emploi, salaires ou autre), ce qui relève de différences légitimes de dotations et ce qui relève de différences de traitement de ces dotations. Par exemple, on va considérer que le type de diplôme, comme dotation spécifique de capital humain pour un individu, va être une cause structurelle légitime de différences d'insertion, et de variations dans les taux de chômage ou les salaires, c'est à dire qu'on accepte qu'il puisse justifier des différences de niveau entre groupes d'individus porteurs de diplômes différents. On accepte donc l'idée que le diplôme est un élément de différenciation justifié sur le marché du travail. Cependant, cette question de la légitimité des différences suppose également que, pour une dotation spécifique donnée, on observe une même valorisation pour tous les individus qui en sont porteurs. Si par exemple on accepte qu'il y ait une différence de salaire du fait d'avoir tel diplôme, on s'attend à ce que l'effet diplôme soit le même pour tous. Quand ce n'est pas le cas, on observe alors des différences de traitement, que l'on va considérer comme injustifiées. La modélisation économétrique va permettre de distinguer, dans l'écart brut, ce qui relève d'une différence injustifiée de ce qui relève d'une différence de dotation.

Une telle approche a ses limites. Comme dans toute modélisation économétrique, elle est fortement dépendante du cadre théorique dans lequel elle s'inscrit et qui définit les facteurs de différenciation légitimes sur le marché du travail. De plus, pour un cadre théorique donné, il y a également nécessité de spécifier le bon modèle : quand on omet une variable de dotation, on modifie les estimations de rendement des autres dotations, et on obtient une estimation biaisée.

Deuxième limite de ce type d'approche : on estime ici une discrimination moyenne, supportée par un groupe, mais à aucun moment on ne peut identifier une discrimination qui ne toucherait qu'une partie des individus d'un groupe donné.

Qu'est-ce que cela donne concrètement sur l'accès à l'emploi ? Les résultats suivants proviennent d'une étude réalisée par des collègues, Arnaud Dupray et Stéphanie Moullet, qui porte sur l'accès à l'emploi des jeunes sortis de formation initiale en 1998 et observés en 2001 [Dupray A. et Moullet S. : *L'insertion des jeunes d'origine maghrébine en France*, Net.Doc N° 6, Céreq, mai 2004].

Etude Dupray-Moulet (*)	
Ecart dans le taux d'accès à l'emploi (2001)	16,3 points
<i>dont :</i>	
lié aux différences de dotations	10,8 points
lié aux différences de traitement	5,5 points
Modélisation prenant en compte des informations sur :	
Diplôme atteint - Emploi ou stages pendant les études - formation par alternance - région de résidence à la fin des études - situation et profession de chacun des parents - situation familiale (enfant - vie en couple) - sexe	

(*) Dupray A. et Moulet S. : L'insertion des jeunes d'origine maghrébine en France, *Net.Doc N°6, Céreq, mai 2004*

Entre les deux groupes étudiés, les jeunes dont les parents sont nés en France et ceux dont les parents sont nés au Maghreb, on constate un écart d'environ 16 points dans l'accès à l'emploi. Une fois pris en compte tout un ensemble de caractéristiques (le diplôme atteint, le fait que l'individu a travaillé pendant ses études, qu'il sort d'une formation par alternance, qu'il résidait en Ile de France au moment de la fin de ses études, son origine sociale, sa situation familiale et son sexe), la modélisation indique que les deux tiers de cet écart s'expliquent par des différences de dotations (10,8%) et que le tiers restant (5,5%) reste imputable à des différences de traitement, considérées comme non justifiables, et qui renverraient donc à de la discrimination. L'étude d'Olivier Joseph et Séverine Lemièrre [*La discrimination de genre et d'origine à l'encontre des jeunes sur le marché du travail. Mesures à partir de différents aspects des situations professionnelles*, *Net.Doc N° 12, Céreq, mars 2005*], basée sur des modélisations et des champs un peu différents, présente des résultats convergents.

Ce type d'analyse peut être étendu à d'autres dimensions du parcours professionnel : les écarts de rémunération ou encore les écarts d'accès au contrat à durée indéterminée. Concernant la discrimination salariale liée à l'origine, on en observe finalement très peu une fois pris en compte le fait qu'une partie seulement des jeunes a accédé à l'emploi. En revanche, on observe une discrimination liée à l'origine dans l'accès au contrat à durée indéterminée, pour les hommes.

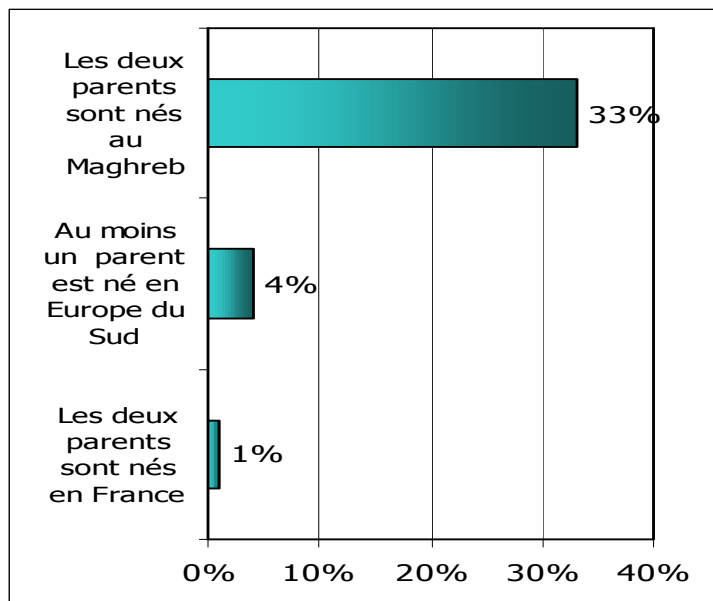
La méthode déclarative

Dans l'enquête « Génération 98 », des questions très directes sont posées aux enquêtés à chacune des interrogations, du type : « Estimez-vous avoir été victime d'une discrimination dans le travail sur la période passée ? ». Si la personne répond par l'affirmative, d'autres questions portant sur les motifs de cette discrimination et sur les formes prises par cette discrimination seront posées.

La méthode déclarative présente aussi des limites. D'abord, elle repose sur une information déclarative, c'est-à-dire sur le ressenti déclaré par l'individu, avec toutes les ambiguïtés que cela comporte. Outre la subjectivité inhérente à une telle perception, il semblerait que, dans certains cas, l'énoncé d'une discrimination renvoie aussi à une forme d'affirmation collective, au sens où les individus affirment une discrimination parce qu'ils ont la volonté de témoigner pour les autres individus du groupe social auquel ils s'identifient. Ensuite, l'individu ne peut rapporter que la discrimination qu'il perçoit. Or toutes les situations où la discrimination est à l'œuvre n'ont pas le même degré de visibilité, et par exemple les discriminations à l'embauche sont difficiles à déceler par les candidats qu'elles touchent.

Qu'est-ce que cela donne au concret sur l'accès à l'emploi ? Dans le tableau suivant, parmi ceux ayant déclaré avoir été victime de discrimination au moins une fois, on a isolé ceux qui ont mentionné le motif de discrimination raciale : 33% des jeunes d'origine maghrébine sont concernés.

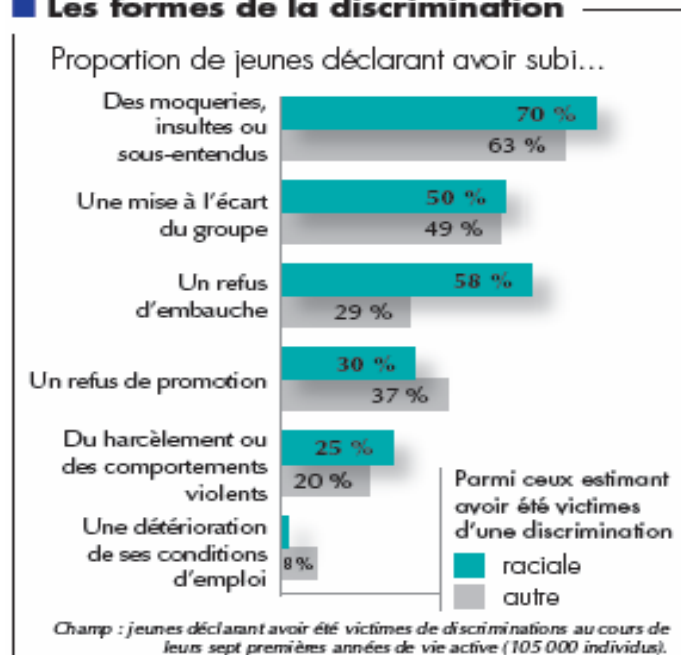
Sentiment de discrimination raciale et origine nationale



Source : Couppié, Gasquet, Lopez (coord.) : Quand la carrière commence..., Céreq, 2007.

Deuxième information recueillie : les formes de discrimination. Parmi les jeunes ayant indiqué le motif de discrimination raciale, 58% déclarent que cela s'est traduit par un refus d'embauche. Si on fait une rapide règle de trois, on a donc presque un jeune sur cinq d'origine maghrébine qui déclare avoir été, à un moment donné de son parcours, victime de discrimination raciale à l'embauche.

■ Les formes de la discrimination



Source : Couppié, Gasquet, Lopez (coord.) : Quand la carrière commence..., Céreq, 2007.

Conclusion

Les résultats des deux approches, économétrique et déclarative, convergent. On a bien, en ce qui concerne l'accès à l'emploi, une double présomption de discrimination à l'encontre des jeunes d'origine maghrébine. Ceci vient s'ajouter à leur désavantage en termes de dotations, désavantage très net en ce qui concerne le niveau de diplôme atteint dans le système éducatif, ce qui est corrélé au fait que ce sont des jeunes qui sont massivement issus des milieux populaires.

Fiche signalétique

Intervenant : **Thomas Couppié, chargé d'études au Deva (Département des entrées dans la vie active) du Céreq (Centre d'études et de recherches sur les qualifications)**

Nom de l'opération : **Enquêtes "Génération 98" du Céreq**

Source(s) des données (enquête spécifique, exploitation d'enquête(s) générale(s), exploitation de fichier(s) de gestion) : Exploitation d'enquêtes générales portant sur l'insertion professionnelle des jeunes sortant de formation initiale. Les résultats présentés ici portent sur la « Génération 98 » interrogée trois, cinq et sept ans après leur sortie de formation initiale, mais d'autres cohortes (1992, 2001, 2004) ont également été interrogées.

Lieu(x) de la collecte : France

Mode(s) de recueil (interne ou externe, auto-administré ou par enquêteur ; téléphone, face à face, internet, postal...) : Enquêtes téléphoniques réalisées par des enquêteurs. Sous-traitance des entretiens à une société extérieure privée.

Population concernée : Individus sortant de formation initiale au cours d'une année donnée.

Type et taille de l'échantillon, ou effectif de la collecte exhaustive : Selon la cohorte et la vague d'enquête, entre 16.000 (Génération 98 à sept ans) et 70.000 questionnaires (Génération 2004) représentatifs d'une cohorte de 750.000 jeunes.

Taux de réponse : 73% (Génération 98 interrogée en 2005).

Donnée(s) utilisée(s) pour construire les catégories d'"origines" : Lieu de naissance de l'enquêté, lieu de naissance des deux parents de l'enquêté, nationalité d'origine des deux parents de l'enquêté.

Principales questions explorées en lien avec les variables d'"origines" : Insertion professionnelle et débuts de carrière des jeunes issus de l'immigration.

Principaux résultats : Les jeunes issus de l'immigration connaissent des conditions d'insertion plus difficiles. Les difficultés se concentrent sur l'accès à l'emploi et se traduisent par un risque de chômage accru et persistant dans le temps, et un accès moindre aux formes d'emploi à durée indéterminée (fonctionnaires, CDI).

Références :

Céreq (2007) : *Quand la carrière commence... Les sept premières années de vie active de la Génération 98*. Editions Céreq, Marseille, 2007. (à commander au Céreq - bon de commande disponible ici : <http://www.cereq.fr/pdf/bondecommandecarriere.pdf>)

Silberman R. et Fournier I. (2006) : "Jeunes issus de l'immigration. Une pénalité à l'embauche qui perdure...", *Bref* n°226, janvier 2006. (téléchargeable ici : <http://www.cereq.fr/pdf/b205.pdf>)

Le numéro 94 (avril-juin 2006) de la Revue *Formation Emploi* consacre un dossier au thème : "Immigration et marché du travail". (sommaire : <http://www.cereq.fr/pdf/fe94.pdf>).

Lopez A. et Thomas G. (2006) : "L'insertion des jeunes sur le marché du travail : le poids des origines socioculturelles", *Données Sociales*, INSEE, 2006.

Frickey A. et Primon J.-L. (2004) : "Les jeunes issus de l'immigration. De l'enseignement supérieur au marché du travail", *Bref* n°205, février 2004. (téléchargeable ici : <http://www.cereq.fr/pdf/b205.pdf>)

Méthodes et résultats de l'étude « Porter un prénom arabe ou musulman est-il discriminant dans l'enseignement supérieur ? »

Marie-Noëlle Decharne et Eric Liedts
Pôle Universitaire de Lille Nord-Pas de Calais

L'Observatoire Régional des Etudes Supérieures est une des missions du Pole Universitaire de Lille. Notre mission est de fournir aux décideurs de la région des outils d'aide à la compréhension des mécanismes à l'œuvre dans l'enseignement supérieur. Nous produisons donc une série d'indicateurs et d'études statistiques pouvant porter sur des populations d'étudiants (ensemble des inscrits, nouveaux bacheliers, étudiants étrangers, etc.), sur des filières particulières (IUT, formations du secteur sanitaire et social, etc....) ou sur des variables particulières (rôle de la variable sexe, de l'origine sociale sur les parcours étudiants). Nous réalisons également des enquêtes régionales (en particulier des enquêtes d'insertion des diplômés) avec la collaboration des observatoires des établissements.

Objet de l'étude et population

L'ORES dispose d'un outil unique en France : un fichier de l'ensemble des inscrits physiques (sans double-compte) dans l'Enseignement Supérieur (ES) public et privé du Nord-Pas de Calais (NPdC), c'est-à-dire à peu près 150.000 étudiants par an. Il est constitué à partir de la collecte annuelle des renseignements fournis par les étudiants au moment de leur inscription. Il permet de connaître et de quantifier toutes leurs caractéristiques, et aussi de suivre leurs parcours dans des analyses longitudinales. C'est donc un outil d'une très grande richesse qui autorise des traitements fins et approfondis.

En mai 2005, nous avons initié une nouvelle série d'études portant sur la question des inégalités dans l'enseignement supérieur. La première variable sur laquelle nous avons fait porter l'analyse a été le genre. En partant d'un constat bien connu, celui de l'orientation différenciée des filles et des garçons dans l'enseignement supérieur, nous avons essayé de comprendre comment ces différences se construisent socialement et comment se décompose la complexité des mécanismes producteurs d'inégalités.

Selon la même démarche, et en nous inspirant d'autres travaux (et plus particulièrement de ceux de G. Felouzis), nous avons proposé d'examiner la situation dans l'ES NPdC d'une population d'étudiants porteurs d'une autre caractéristique discriminante : le prénom.

Le prénom est utilisé ici comme l'indicateur d'une identité sociale que l'on peut qualifier de «stigmatisable». Selon Erving Goffman, un individu est dit stigmatisé lorsqu'il présente un attribut qui le disqualifie lors de ses interactions avec autrui. Le prénom est considéré ici comme un facteur de discrimination sociale. Le prénom n'est pas conçu au sens strict comme un indicateur de l'origine ethnique. Les débats autour de ce concept sont suffisamment anciens et riches pour qu'il soit admis que le prénom n'en est qu'un indicateur, plus ou moins fiable, parmi de nombreux autres.

La population à propos de laquelle nous menons l'étude des parcours est celle des étudiants de nationalité française. En effet, il nous a semblé inopportun d'inclure dans le champ de l'étude

la population des étudiants de nationalité étrangère : nos études montrent qu'elle présente des caractéristiques sociodémographiques et des parcours dans l'enseignement supérieur du Nord-Pas de Calais bien différents de ceux des étudiants français. De plus, introduire la nationalité pourrait conduire à brouiller les effets redevables de la seule variable liée à « l'identité stigmatisable ».

La méthode des prénoms

Les prénoms des étudiants de nationalité française, recensés dans le fichier régional, ont été utilisés pour créer deux catégories de prénoms :

- celle des prénoms d'origine arabe ou musulmane
- celle des autres prénoms, quelle que soit leur origine

A partir du fichier de gestion (fichier nominatif des étudiants inscrits dans la région NPdC), nous avons extrait la liste (anonyme) de tous les prénoms portés par les étudiants. A chacun des prénoms a été associé le caractère « prénom arabe ou musulman » ou « autre prénom ». Seule cette variable a été réintégrée dans le fichier des inscrits, duquel ont été supprimés les noms et les prénoms, pour aboutir au fichier de travail anonymisé qui a été utilisé pour les traitements statistiques. A aucun moment de la préparation du fichier, la catégorie de prénom n'a donc été rapprochée de l'identité des étudiants et a fortiori réintégrée dans le fichier de gestion.

L'expertise sur les prénoms a été réalisée par une association lilloise (Association pour une Interculturalité Citoyenne), auteure de « L'indispensable manuel orthographique des prénoms français d'origine arabe ou musulmane » (A. Chebira, *L'indispensable manuel orthographique des prénoms français d'origine arabe ou musulmane*. Editions APIC, 2005). A partir du recensement, réalisé annuellement par l'INSEE, des prénoms qui sont donnés à l'état civil français, cet ouvrage identifie les prénoms d'origine arabe ou musulmane ainsi que toutes leurs orthographes répertoriées en France ; et c'est ce qui en fait son grand intérêt pour nous.

Les 149 102 inscrits en 2004/2005 dans la totalité de l'ES NPdC, portent 17 814 prénoms (ou formes orthographiques) différents. Parmi eux, 3 259 ont été identifiés comme étant d'origine arabe ou musulmane. Ont été exclus de cette catégorie ceux qui, malgré leur origine arabe ou musulmane, présentent une forte proximité avec des prénoms français ou internationaux (Myriam, Nadia, Medhi, etc.).

Sur la population de référence de l'étude, constituée des nouveaux bacheliers 2004, de nationalité française inscrits dans l'ES NPdC en 2004-2005, 2 059 personnes ont été repérées comme porteuses d'un prénom d'origine arabe ou musulmane, et 25 459 ont été classées dans la catégorie des "autres prénoms".

Problématique

L'objectif de l'étude n'est évidemment pas de réaliser un simple "comptage" des étudiants portant un prénom arabe ou musulman. Elle vise à mesurer l'influence respective d'un certain nombre de variables dont la sociologie de l'éducation a déjà montré le rôle discriminant sur

les cursus dans l'enseignement supérieur : sexe, origine sociale, type de baccalauréat, précocité au bac.

Elle met en œuvre des outils statistiques qui permettent la comparaison systématique des deux populations d'étudiants : les "porteurs d'un prénom arabe ou musulman" et les "autres".

Pour répondre à la question : « porter un prénom arabe ou musulman est-il discriminant dans l'enseignement supérieur ? », nous nous sommes intéressés plus particulièrement à deux étapes du cursus étudiant :

- la première inscription des nouveaux bacheliers dans les filières d'études de l'ES
- leur devenir après un an d'études n+1 (2005/2006) : abandon, redoublement, passage à un niveau supérieur.

A chacune de ces deux étapes, le rôle du prénom est systématiquement comparé à celui des autres variables :

- sexe,
- origine sociale,
- type de baccalauréat,
- précocité au bac.

Démarche statistique

Les porteurs d'un type de prénom ou d'un autre ont-ils des « sorts » différents dans l'enseignement supérieur (orientation, devenir...) ?

1. Dans un premier temps, on se demande si les différences ne sont pas dues aux caractéristiques de ces populations au regard d'autres variables explicatives

En effet, on constate en phase exploratoire que les porteurs d'un prénom arabe ou musulman ont des caractéristiques différentes des autres : ce sont plus souvent des filles, ils sont plus âgés en moyenne, plus souvent en retard à l'obtention du bac, plus souvent originaires de catégories sociales modestes, moins souvent titulaires d'un bac général. Pour décomposer les mécanismes explicatifs, on associe les variables 2 à 2 pour vérifier la prépondérance de l'une sur l'autre.

Par exemple, ayant constaté, par de simples tris croisés, que :

- les étudiants porteurs d'un prénom arabe ou musulman inscrits dans l'enseignement supérieur sont moins souvent que les autres titulaires d'un bac général,
- et que, par ailleurs, les bacheliers généraux s'inscrivent plus souvent dans des filières prestigieuses,

on associe les variables « type de bac » et « prénom » pour estimer laquelle des deux variables influence le plus l'orientation dans une filière d'études.

Effectif % colonne + -	Bac Général		Bac Techno		Bac Pro		TOTAL
	prénom arabe / musulman	autre prénom	prénom arabe / musulman	autre prénom	prénom arabe / musulman	autre prénom	
DEG-AES	171 19% +++	2116 13% +++	218 26% +++	538 7% ---	149 42% +++	171 13% ---	3363 12%
LAL-SHS	207 24% +++	3932 24% +++	153 18% ---	581 8% ---	115 33% +++	89 7% ---	5077 18%
Sciences	113 13% ++	2272 14% +++	29 3% ---	379 5% ---	10 3% ---	52 4% ---	2855 10%
Santé	131 15% +++	1578 10% +++	7 1% ---	31 0% ---	2 1% ---	---	1749 6%
IUT	91 10%	1627 10%	92 11%	883 11% +++	4 1% ---	60 5% ---	2757 10%
CPGE Ec	85 10%	2671 16% +++	6 1% ---	71 1% ---	---	---	2833 10%
STS	61 7% ---	1285 8% ---	309 37% +++	4608 60% +++	70 20% ---	864 68% +++	7197 26%
Autre Filière	18 2% ---	987 6%	17 2% ---	621 8% +++	1 0% ---	43 3% ---	1687 6%
TOTAL	877 100%	16468 100%	831 100%	7712 100%	351 100%	1279 100%	27518 100%

Lecture du tableau

Description des cellules :

- En haut, à droite : les effectifs
- En bas à gauche : les % en colonne
- En bas à droite : le signe des associations locales. S'il est positif, il indique une surreprésentation, s'il est négatif il indique une sous-représentation. Plus le nombre de signes est important, plus les écarts observés sont significatifs.

Les cellules bleues sont celles où les inscriptions sont significativement les plus fréquentes, et les cellules roses où elles sont significativement les moins fréquentes.

On voit nettement que les principales différences sont dues au type de bac, et ceci quelle que soit la catégorie de prénom :

- les bacheliers généraux sont plus fréquemment inscrits dans les filières universitaires générales et de Santé ainsi qu'en IUT et CPGE (classes préparatoires aux grandes écoles). Ils sont notablement moins présents en STS ;
- les bacheliers technologiques, sont souvent inscrits en IUT et STS et moins ailleurs.

Ce n'est que dans un deuxième temps qu'on constate des différences selon le prénom dans chacun des types de bac. Ainsi, il ressort très nettement que moins le type de bac est valorisé dans l'enseignement supérieur et plus les différences d'orientation entre les étudiants ayant un prénom arabe ou musulman et les autres sont fortes. Par exemple, chez les bacheliers professionnels, avoir un prénom arabe ou musulman augmente très significativement les chances d'être inscrit à l'université en DEG-AES et en LAL-SHS et diminue très fortement celle de l'être en STS.

On mène ainsi l'analyse en associant toutes les variables 2 à 2.

2. Ensuite, on cherche à évaluer le rôle de toutes les variables en même temps et des effets de masse qu'elles produisent. On procède à une approche multidimensionnelle, « toutes choses inégales réunies »

Toutes les modalités des variables introduites dans l'analyse sont projetées simultanément sur un même plan (dit factoriel) en fonction de leur proximité entre elles et aussi des masses (effectifs) qu'elles représentent.

Dans notre exemple, on cherche à expliquer l'orientation des bacheliers dans les filières d'études. Le plan factoriel que nous avons obtenu est reproduit page suivante.

L'axe horizontal, premier axe factoriel, oppose deux groupes de filières :

- à gauche, les filières d'études considérées comme étant les plus prestigieuses (CPGE, Santé et Sciences) ;
- à droite, les STS.

Comment se projettent sur le plan factoriel les variables explicatives de cette orientation ?

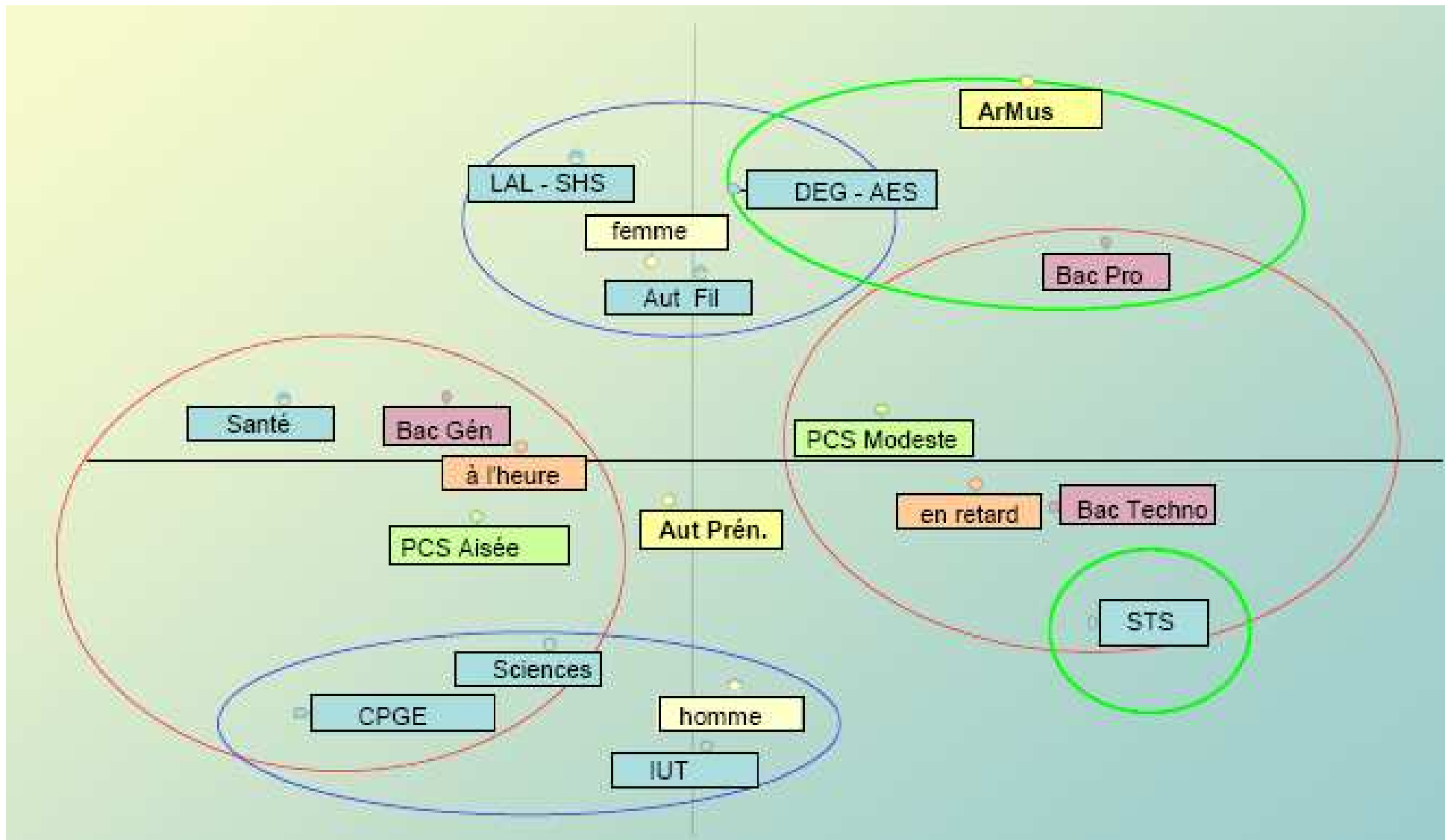
- les filières les plus prestigieuses (à gauche) sont proches des origines sociales aisées, des séries de bac générales, et du fait d'être "à l'heure" au bac.
- à l'opposé, les STS voisinent avec les origines modestes, les séries technologiques et professionnelles du bac, le retard scolaire.

Ce premier axe montre la prédominance de "l'origine sociale", de la "précocité" et du "type de bac" sur l'orientation dans l'enseignement supérieur.

Mais ces variables n'expliquent pas toutes les différences. Le deuxième axe, vertical, oppose également deux groupes de filières :

- en haut, LAL-SHS, DEG-AES, la catégorie des Autres filières ;
- en bas les Sciences, les CPGE et les IUT.

Le sexe est la première variable qui contribue à cette opposition : les femmes sont associées au premier groupe, les hommes au second. Les différences d'orientation expliquées par le prénom n'apparaissent qu'en troisième lieu. Les porteurs d'un prénom arabe ou musulman sont à proximité des filières de DEG-AES et éloignés des STS.



3. Dans un 3ème temps, on neutralise les autres différences qui caractérisent les populations, pour vérifier s'il persiste des différences significatives qui ne s'expliqueraient que par le prénom, dans une approche « toutes choses égales par ailleurs »

On veut évaluer, par exemple, l'influence de chacune des variables sur la probabilité d'être titulaire d'un bac général plutôt que d'un autre type de bac, quand on est inscrit dans l'enseignement supérieur.

On se base sur la probabilité d'une personne de référence définie comme ayant les caractéristiques les plus fréquentes dans la population : cette personne de référence est une femme, d'origine sociale aisée, à l'heure au bac, portant un prénom autre qu'arabe ou musulman. Elle a 88 % de chances d'être titulaire d'un bac général (plutôt qu'un autre type de bac) quand elle s'inscrit dans l'ES NPdC.

Parmi les inscrits dans l'enseignement supérieur	
	Effets marginal
▪ Une femme	=> - 5 points (homme)
▪ D'origine aisée	=> - 17 points (modeste)
▪ A l'heure	=> - 24 points (en retard)
▪ Ne portant pas un prénom arabe ou musulman	=> - 4 points (prénom arabe)
a 88% de chances d'être ...	a x % de chances d'être ...
... titulaire d'un baccalauréat général	

On regarde ensuite comment varie cette probabilité si on change, une à une, les caractéristiques :

- pour un homme, la probabilité d'avoir un bac général (toutes les autres caractéristiques restant identiques) diminue de 5 points. Elle passe à 83 % ;
- pour une étudiante d'origine sociale modeste, elle diminue de 17 points. Elle passe à 71 % ;
- pour une étudiante en retard (au moment de l'obtention du bac), elle diminue de 23 points. Elle passe à 65 % ;
- pour une étudiante portant un prénom arabe ou musulman, elle baisse de 4 points. Elle passe à 84 %.

Ainsi, peut-on constater que les deux variables qui influencent le plus la probabilité d'avoir un bac général sont d'abord la précocité puis l'origine sociale.

On a également procédé à une analyse « toutes choses égales par ailleurs » appliquée à l'orientation : pour chaque filière, la probabilité de s'y inscrire est comparée à celle de la personne de référence que nous avons décrite ci-dessus. Le traitement permet donc de vérifier si le fait de porter un prénom arabe ou musulman plutôt qu'un autre joue sur la probabilité de s'inscrire dans une filière, toutes choses égales par ailleurs (c'est-à-dire toutes les autres caractéristiques connues introduites dans l'analyse).

Filières universitaires :

Filière d'études	DEG-AES	LAL-SHS	Sciences	Santé
Variables	Effet marginal de la modalité	Effet marginal de la modalité	Effet marginal de la modalité	Effet marginal de la modalité
PCS modeste	+1,4 pt	+1,9 pt	+1,7 pt	-5,7 pt
PCS aisée	réf	réf	réf	réf
Homme	-3,6 pt	-16,5 pt	+8,2 pt	-2,9 pt
Femme	réf	réf	réf	réf
Prénom Arabe ou Musulman	+15,7 pt	+10,8 pt	-1,7 pt	+13,7 pt
Autre Prénom	réf	réf	réf	réf
Bac Pro	+2,3 pt	-15,0 pt	-7,0 pt	-14,1 pt
Bac Techno	-5,2 pt	-18,4 pt	-6,5 pt	-13,4 pt
Bac Gén	réf	réf	réf	réf
en retard	+2,3 pt	+5,5 pt	+1,6 pt	-7,3 pt
à l'heure	réf	réf	réf	réf
Probabilité d'être inscrit dans la filière	12,7%	28,3%	8,9%	14,4%

Filières sélectives :

Filière d'études	IUT	CPGE	STS	Aut Fil
Variables	Effet marginal de la modalité	Effet marginal de la modalité	Effet marginal de la modalité	Effet marginal de la modalité
PCS modeste	-0,5 pt	-7,4 pt	+3,0 pt	+0,1 pt
PCS aisée	réf	réf	réf	réf
Homme	+7,2 pt	+20,2 pt	+1,3 pt	-4,0 pt
Femme	réf	réf	réf	réf
Prénom Arabe ou Musulman	-0,1 pt	-2,0 pt	-3,9 pt	-4,3 pt
Autre Prénom	réf	réf	réf	réf
Bac Pro	-4,5 pt	-15,9 pt	+43,2 pt	-2,7 pt
Bac Techno	+0,3 pt	-14,5 pt	+41,7 pt	+2,0 pt
Bac Gén	réf	réf	réf	réf
en retard	+1,0 pt	-12,0 pt	+0,9 pt	+2,1 pt
à l'heure	réf	réf	réf	réf
Probabilité d'être inscrit dans la filière	6,9%	15,9%	5,8%	7,7%

Dans le groupe des filières universitaires :

La probabilité, pour notre personne de référence (une femme, d'origine sociale aisée, n'ayant pas un prénom arabe ou musulman, titulaire d'un bac général obtenu à l'heure), d'être inscrite en DEG-AES, est de 12,7 % (28,3 % en LAL-SHS, 8,9 % en Sciences, 14,4 % en Santé).

Le fait de porter un prénom arabe ou musulman est la caractéristique qui augmente le plus les chances d'être inscrit dans tous les groupes de disciplines universitaires sauf en Sciences. Ainsi, par exemple, en DEG-AES, changer le type de prénom (en gardant les autres caractéristiques identiques) augmente la probabilité d'inscription dans la filière de 15,7 pts, la faisant passer ainsi à 28,4 %.

Dans le groupe des filières sélectives :

Le fait de porter un prénom arabe ou musulman n'augmente ni ne diminue les chances d'être inscrit en IUT ou en CPGE ; par contre, il réduit fortement celles d'être inscrit en STS : 3,9 points sur une probabilité de référence de 5,8 % ; ce qui la divise par 3.

On en tire la conclusion que dans les filières sélectives, le prénom n'apparaît pas comme la variable la plus discriminante, c'est plus souvent la série de bac (par exemple en STS), voire le sexe (par exemple en CPGE).

Conclusion

Nous avons bien sûr procédé au même type d'analyse à propos du devenir des étudiants un an après leur première inscription. Nous avons évalué le rôle du prénom, toutes choses égales par ailleurs, sur la probabilité d'abandonner puis sur celle de redoubler.

En résumé, l'étude prouve que porter un prénom d'origine arabe ou musulmane a bien des effets discriminants sur l'orientation et la poursuite d'études dans l'enseignement supérieur. Toutefois, ces effets s'associent à ceux produits par d'autres caractéristiques attachées à cette population spécifique d'étudiants : la série de baccalauréat, la précocité au baccalauréat, l'origine sociale et le sexe. S'ils n'ont pas les mêmes chances d'accéder aux mêmes filières de l'enseignement supérieur que les autres, ni d'y réussir, c'est aussi qu'ils sont moins bien lotis que les autres, tant socialement que scolairement.

Vous pouvez vous reporter à l'étude pour en découvrir les principaux résultats.
http://www.poleuniv-lille-npdc.fr/n_telecharg/ores/04-05_B_R001.pdf

Fiche signalétique

Intervenants : **Marie-Noëlle Decharne et Eric Liedts (ORES – Pôle Universitaire de Lille NPdC)**

Nom de l'opération : « **Porter un prénom arabe ou musulman est-il discriminant dans l'enseignement supérieur ?** »

Source(s) des données (enquête spécifique, exploitation d'enquête(s) générale(s), exploitation de fichier(s) de gestion) : Exploitation du fichier régional de l'ensemble des inscrits dans l'ensemble de l'enseignement supérieur (public et privé) de la région Nord – Pas de Calais, expurgé des doubles comptes (à peu près 150 000 étudiants par an). Le fichier existe depuis l'année universitaire 1992-1993. Ce fichier nominatif fait l'objet d'une déclaration à la CNIL.

Mode(s) de recueil (interne ou externe, auto-administré ou par enquêteur ; téléphone, face à face, internet, postal...): Le fichier régional est constitué par la collecte annuelle des données d'inscription des 250 établissements d'enseignement supérieur de la région.

Population ou périmètre concernés : La population concernée par l'étude est celle des bacheliers de l'année 2004 inscrits dans l'enseignement supérieur du Nord – Pas de Calais (ES NPdC) en 2004-2005

Type et taille de l'échantillon, ou effectif de la collecte exhaustive : 24 289 bacheliers 2004 se sont inscrits dans l'ES NPdC en 2004-2005. La population a été répartie en deux catégories sur le critère du prénom :

- les bacheliers qui portent un prénom arabe ou musulman : 1 936
- les autres bacheliers : 22 353

Taux de réponse : collecte exhaustive

Donnée(s) utilisée(s) pour construire les catégories d'"origines" : le prénom. L'utilisation du prénom a fait l'objet d'une demande d'autorisation puis de nombreuses discussions avec la CNIL.

Libellés de ces catégories :

- Nouveaux bacheliers porteurs d'un prénom arabe ou musulman
- Autres nouveaux bacheliers

Principales questions explorées en lien avec les variables d'"origines" : Le rôle de la variable « porter ou non un prénom arabe ou musulman » sur l'orientation des bacheliers et sur leur devenir un an après est systématiquement comparé à celui d'autres variables utilisées traditionnellement en sociologie de l'éducation : sexe, origine sociale, série du baccalauréat, précocité au baccalauréat.

L'orientation est décrite par la filière d'inscription dans l'enseignement supérieur.

Le devenir un an après fait l'objet de trois catégories : la sortie de l'enseignement supérieur régional, le redoublement, et le passage à bac+2.

Deux types de tests statistiques ont été utilisés :

- l'analyse factorielle de correspondances multiples permet de représenter simultanément l'effet de toutes les modalités des variables (sur l'orientation puis sur le devenir)
- la régression logistique multinomiale permet d'estimer l'effet de chacune des variables explicatives « toutes choses égales par ailleurs » sur les probabilités d'orientation puis de devenir

Résumé des résultats :

L'étude prouve que porter un prénom d'origine arabe ou musulmane a bien des effets discriminants sur l'orientation et la poursuite d'études dans l'enseignement supérieur. Toutefois, ces effets s'associent à ceux produits par d'autres caractéristiques attachées à cette population spécifique d'étudiants : la série de baccalauréat, la précocité au baccalauréat, l'origine sociale et le sexe. S'ils n'ont pas les mêmes chances d'accéder aux mêmes filières de l'enseignement supérieur que les autres, ni d'y réussir, c'est aussi qu'ils sont moins bien lotis que les autres, tant socialement que scolairement.

Résultats détaillés :

- La première étape est celle de la description :

Les primo-entrants dans l'enseignement supérieur portant un prénom arabe ou musulman se distinguent des autres :

- ils sont légèrement plus souvent des filles
- ils sont plus âgés et aussi plus souvent en retard à l'obtention du baccalauréat
- ils sont plus souvent originaires de catégories sociales défavorisées
- ils ont moins souvent décroché une mention au baccalauréat
- ils sont moins souvent titulaires d'un bac général, mais plus souvent d'un bac technologique ou professionnel

- Dans quelles filières de l'ES NPdC les nouveaux bacheliers se sont-ils inscrits en 2004/2005 ?

Les nouveaux bacheliers portant un prénom arabe ou musulman s'inscrivent plus souvent que les autres dans les filières générales de l'université, en particulier dans deux groupes de disciplines (Droit-Economie-Gestion-AES et Lettres-Arts-langues-Sciences Humaines et Sociales). En contrepartie, ils s'orientent globalement moins souvent que les autres dans les filières sélectives.

Cependant, quand on tient compte simultanément de toutes les variables, on observe que l'orientation dans l'ES NPdC ne dépend pas prioritairement du prénom porté. C'est plutôt le type de bac qui la conditionne. Mais, moins le type de bac est valorisé dans l'enseignement supérieur (*ex : bac professionnel*) et plus les différences d'orientation entre les étudiants ayant un prénom arabe ou musulman et les autres sont fortes.

Si on examine isolément l'effet de chacune des variables considérées, le prénom apparaît de manière très significative comme produisant les discriminations les plus fortes. Porter un prénom arabe ou musulman plutôt qu'un autre favorise l'inscription dans toutes les filières générales universitaires sauf en Sciences. Dans les filières sélectives, le fait de porter un prénom arabe ou musulman n'a quasiment pas d'influence (*toutes choses égales par ailleurs*) sur celui de s'inscrire en IUT ou en Classe Préparatoire aux Grandes Ecoles (*plutôt que de ne pas s'y inscrire*) ; par contre, il réduit considérablement les chances de s'inscrire en Section de Techniciens Supérieurs.

- Quel est le devenir en 2005/2006 des nouveaux bacheliers inscrits en 2004/2005 ?

La comparaison des trois situations possibles dans chacune des deux populations révèle que la sortie de l'enseignement supérieur régional après un an d'études (15 % des étudiants) n'est pas sensible au prénom. Par contre, le redoublement est deux fois plus fréquent chez les porteurs d'un prénom arabe ou musulman que chez les autres (50 % contre 25 %), bien évidemment au détriment du passage à un niveau d'études supérieur qui suit la tendance inverse. Prénom, type de bac et filières d'études combinent leurs effets pour aboutir à ces situations.

Quand on observe simultanément toutes les variables, le clivage observé dans les filières universitaires générales entre passage et "sortie/redoublement" n'est pas expliqué par le prénom, mais par l'origine sociale, la précocité au bac, et le type de bac.

Cependant, toutes choses égales par ailleurs, le fait d'avoir un prénom arabe ou musulman a un effet propre. Dans les filières universitaires générales, il diminue le risque de quitter l'ES-NPdC et augmente considérablement le fait de "redoubler". Dans les filières sélectives, s'il ne produit pas d'effet sur le fait de quitter l'ES NPdC après une année d'études, il augmente également très fortement le risque de se réinscrire à un niveau bac+1.

Référence :

« Porter un prénom arabe ou musulman est-il discriminant dans l'enseignement supérieur ? Orientation et poursuite d'études dans l'enseignement supérieur du Nord – Pas de Calais. »

Marie-Noëlle Decharne et Eric Liedts – ORES Population - Mai 2007

Pôle Universitaire Lille Nord – Pas de Calais

http://www.poleuniv-lille-npdc.fr/n_telecharg/ores/04-05_B_R001.pdf

Études à l'intérieur des organisations

L'enquête expérimentale « Mesure de la diversité » : test de différentes catégorisations des origines et réaction des enquêtés

Patrick Simon, INED

Mon intervention portera sur les catégories. Je me situerai donc en amont de la question de la mesure : il s'agit moins de savoir quelle méthodologie précise peut être utilisée pour mesurer, que de savoir *qui entre* dans la mesure.

Je voudrais présenter, pour cela, le résultat d'une enquête réalisée en juin 2006 – « Mesure de la diversité » –, qui cherchait, premièrement à tester différents types d'enregistrement des origines, et deuxièmement à recueillir les réactions des enquêtés à ces différentes formes de déclaration de l'origine. On se situe donc là au cœur du contentieux et des controverses qui se développent ces dernières années sur les modes d'enregistrement, sur les catégories elles-mêmes et sur la terminologie utilisée pour décrire les personnes exposées aux discriminations.

Les données du problème

Comme Benoît Riandey l'a rappelé, on rencontre un problème récurrent depuis des années tenant à l'insuffisance des données permettant d'observer et d'analyser les discriminations. Ce problème est lié à l'appareil statistique dans son ensemble et renforcé par le flou concernant les populations qui seraient concernées par les discriminations. Il existe des données qui nous permettent de travailler sur des populations étrangères (mais est-ce bien de ces populations dont on veut parler ?), sur les populations immigrées (mais là encore, est-ce bien de ces populations dont on veut parler ?). Au-delà de ces populations, les données disponibles sont extrêmement rares, ce qui invite à réfléchir au développement, dans l'appareil de collecte de données, d'enquêtes faisant référence à d'autres populations que les étrangers ou les immigrés.

Le recueil de données faisant apparaître d'autres catégories que les étrangers ou les immigrés suscite une controverse dite des statistiques « ethniques ». Cette controverse prend de l'ampleur et se durcit ces derniers temps, avec des initiatives gouvernementales qui posent question et, pour ce qui concerne les institutions statistiques (INSEE, CNIL, CNIS) peu d'expérience sur le sujet, si on la compare avec d'autres pays voisins ou plus lointains qui sont également confrontés à une expérience très ancienne d'immigration, de colonisation ou d'esclavage.

Cette controverse s'appuie en particulier sur une hypothèse toujours affirmée mais jamais démontrée : tout le monde serait très opposé à répondre à des questions faisant référence à ses origines, et en particulier les personnes pour qui cette question devient extrêmement sensible parce qu'elles sont elles-mêmes personnellement exposées aux discriminations. Pour le dire autrement, les chercheurs travaillant sur des données quantitatives se font renvoyer une sorte d'irresponsabilité sociale : « vous vous posez des questions, mais les gens ne voudront pas y répondre ».

Protagonistes essentiels du débat, les « gens » ne s'expriment pas vraiment sur la question et jouent le rôle du tiers absent à qui chaque acteur entend faire tenir la position qui renforce son propre argumentaire. La circularité du registre de justification est redoutable : on ne parle pas

d'origine dans les statistiques car cela n'est pas recevable, et ce qui prouve que cela n'est pas recevable est justement que l'on en parle pas (« contraire à la tradition républicaine » dans l'énoncé habituel). La force du sens commun et de la doxa institutionnelle conjugués interdisait de poser la question de l'origine des personnes. On a donc voulu tester la recevabilité de ces catégories, pour situer quel serait le degré de réticence, ou de résistance à l'égard de questions *dans l'éventualité où elles devaient être posées*.

Ce dernier point appelle un commentaire, puisque le cadre dans lequel de telles questions sont formulées déterminent fortement leur acceptabilité. En effet, ce n'est pas la même chose de collecter des données sur l'origine pour étudier les pratiques sociales ou le fonctionnement des systèmes de discrimination ou pour mettre en place des politiques d'exclusion à base raciale ou ethnique, ou démontrer l'infériorité, l'arriération ou plus généralement disqualifier des « groupes ethniques et raciaux ». Le projet qui sous-tend la collecte statistique est autant déterminant que le format des informations collectées et leurs usages ultérieurs. Or au moment de notre enquête, l'utilisation d'informations statistiques à des fins de lutte contre les discriminations est relativement théorique par absence de pratiques antérieures. L'expérience sur laquelle les enquêtés peuvent s'appuyer pour évaluer les risques recelés par de telles données est marquée par les pratiques d'exclusion, voire d'élimination, dans le cadre de la politique antisémite et raciale de Vichy ou de hiérarchisation raciale dans le cadre de la politique coloniale. Il est certain que le *cadrage* n'est pas favorable...

Le cadre de l'enquête

L'enquête a été discutée dans le cadre d'un groupe de travail réunissant des responsables DRH d'entreprises, la situant dans une perspective de recherche action, sachant que par ailleurs, il n'était pas question de mettre en place une méthode ou un monitoring adapté à chacune des entreprises participantes. Il s'agissait avant tout de nourrir la réflexion en amont et d'irriguer le débat plutôt que de produire un outil immédiatement opérationnel. Des « focus groups » nous ont permis de tester le questionnaire et de faire émerger les questions à l'intérieur des entreprises, avec le constat de fortes disparités d'approches selon les entreprises.

Passation des groupes de discussion		
Etablissements	Nombre de questionnaires	Dates
SNCF	12	30/06/2005
RANDSTAD	12	26/10/2005
EAU DE PARIS	14	03/11/2005
AXA	12	15/11/2005
Total	50	

Le problème de l'échantillon, représentatif ou non, était patent dans notre cas. Nous n'avions aucun point de comparaison pour pouvoir faire un échantillon représentatif puisque les questions que nous allions poser ne l'avaient jamais été. On a donc privilégié une approche fondée sur des échantillons raisonnés, c'est-à-dire en essayant de trouver des personnes de caractéristiques extrêmement variées, avec une surreprésentation assez forte des personnes supposées pouvoir répondre à des questions de visibilité face aux discriminations. Une longue périphrase pour nommer des personnes qui seraient susceptibles de remplir des catégories minoritaires.

Comment constituer ces échantillons sans choisir des individus en particulier, donc sans inférer au départ qu'ils sont ou non minoritaires, et sans pouvoir s'appuyer sur des données spécifiques ? Nous avons plutôt réfléchi à partir d'un bassin d'activité, en analysant, avec les entreprises, leurs sites de production et leurs activités susceptibles de concentrer une diversité d'origines, là où d'autres équipes ou ateliers étaient marqués par une très forte homogénéité non minoritaire des membres du personnel. Par ailleurs, les entreprises participantes se différencient également sous l'angle de la diversité interne : on savait par avance que le personnel travaillant aux entrepôts de ED serait plus « minoritaire » que celui des services administratifs d'AXA. De même, nous avons plutôt contacté des universités en périphérie que dans le centre de Paris. Au total, ce sont 1327 questionnaires qui ont été collectés.

Questionnaires collectés	
Etablissements	Nombre de questionnaires
L'OREAL	337
ADECCO	169
ED (GROUPE CARREFOUR)	68
SNCF	219
RANDSTAD	155
AXA	111
EAU DE PARIS	71
UNIVERSITE PARIS 8	42
UNIVERSITE EVRY VAL ESSONNE	61
UNIVERSITE VILLETANEUSE	24
IUFM PARIS	33
CONSEIL REGIONAL IDF	37
Total	1327

Les catégories

Trois approches ont été proposées pour enregistrer l'origine :

- Le pays de naissance, la nationalité, la nationalité des parents, la nationalité des grands parents

L'origine est saisie, comme on le fait de façon ordinaire dans le recensement, à travers le pays de naissance, la nationalité de l'individu, celle de ses parents – ce qui n'est pas demandé dans le recensement - celle de ses grand parents – ce qui n'est pas non plus demandé dans le recensement ni dans les quelques enquêtes qui posent la question des parents. L'enquête porte donc sur deux niveaux d'ascendance.

- Origines géographiques large et fine

Il était demandé aux répondants de « définir » leur origine en se rangeant dans des rubriques renvoyant à des origines géographiques larges (Maghreb, Afrique, Europe, Asie, France), ou des catégories beaucoup plus précises de pays (Algérie, Maroc, Pologne, Italie, Mali, France, etc.).

- Origine « ethno-raciale »

Cette classification qualifiée a posteriori d'ethno-raciale proposait des items comme « blanc », « noir », « arabe ou berbère », « asiatique » et « du sous continent indien ». La liste elle-même, toujours susceptible de discussion, a fait l'objet de révisions successives.

Les deux dernières catégorisations étaient proposées avec des réponses multiples.

A partir de ce matériau, nous avons construit des catégories. Le problème des questions à réponses multiples tient dans le grand nombre de combinaisons possibles pour leur traitement. On traite en effet six ascendants – deux parents et quatre grands parents –, et les combinaisons entre ces six ascendants sont extrêmement importantes. Il faut donc faire des choix. Nous avons opté pour une catégorie incluant les personnes qui ont au moins un ascendant né ou d'origine maghrébine, d'Europe du sud, d'Afrique sub saharienne ou des DOM. De même pour les deux autres catégorisations où les combinaisons sont a priori trop nombreuses pour figurer dans des analyses bivariées. Cependant, les résultats montrent que l'éventail réel des combinaisons est beaucoup plus restreint que le potentiel est qu'il est en partie possible de les reproduire.

La corrélation des catégories

L'une des premières choses que nous voulions tester était de voir en quoi le pays de naissance des parents a une certaine forme de corrélation ou vaut pour une information sur l'origine. Est-ce que, quand on pose la question sur le pays de naissance des parents, on attrape à peu près bien ce que diraient les mêmes personnes quand on leur demande : « quelle est votre origine » ?

Correspondance entre l'ascendance et l'origine déclarée						
Au moins un ascendant	Ont déclaré au moins une origine...					
	Africaine	Antillaise-Caribéenne	Maghrébine	Asiatique	Européenne	Française
Maghreb	9.8	1.2	85.1	0	10.3	34.5
Europe du Sud	0.5	1	2.6	1	75.3	46.9
Afrique Sub Saharienne	85.1	2.7	1.4	1.4	5.4	20.3
DOM	18.2	63.6	1.8	7.3	7.3	36.4

Lire : Les personnes ayant au moins un ascendant originaire du Maghreb ont déclaré une origine « Africaine » dans 9.8% des cas, « Maghrébine » dans 85.1% des cas et « Française » dans 34.5% des cas

85% des personnes qui ont au moins un ascendant d'origine sub-saharienne se sont déclarées d'origine africaine, 63% des personnes qui ont au moins un ascendant originaire des DOM se sont dites d'origine antillaise ou caribéenne, 85% de celles qui ont au moins un ascendant originaire du Maghreb se sont dit d'origine maghrébine. A l'inverse, on peut dire que 15% des personnes qui ont un ascendant qui vient du Maghreb ne se sont pas considéré, ou n'ont pas voulu choisir la catégorie « origine maghrébine ». Cela signifie que quand on prend le pays de naissance des parents, on décrit à peu près 85% des personnes susceptibles de se déclarer d'origine maghrébine. Le recouvrement est un peu moins fort dans le cas des personnes originaires de l'Europe du Sud.

Nous avons posé cette autre question : « est-ce que vous pensez que votre nom évoque une origine étrangère » ? Nous avons donc là le résultat de l'auto perception de cette origine étrangère par les enquêtés. Nous pouvons constater que les personnes qui se déclarent noires considèrent que leur nom a une consonance d'origine étrangère, sans qu'on puisse d'ailleurs vérifier si c'est vrai ou pas, dans 57% des cas. Donc dans 43% des cas, leur nom, d'après eux ou elles, ne les signale pas spécifiquement à l'attention des autres. Par contre, du côté des

personnes qui se disent d'origine arabe ou berbère, 91% d'entre elles considèrent que leur nom les identifie.

Ascendance	
Maghrébine	87.9
Europe du Sud	62.4
Afrique Sub Saharienne	82.7
DOM	25.5
Origine par régions	
Africaine	77.1
Caribéenne	21.7
Maghrébine	87.0
Catégories « ethno-raciales »	
Blanc	29.8
Noir	56.9
Arabe ou Berbère	91.3
Ensemble	40.8

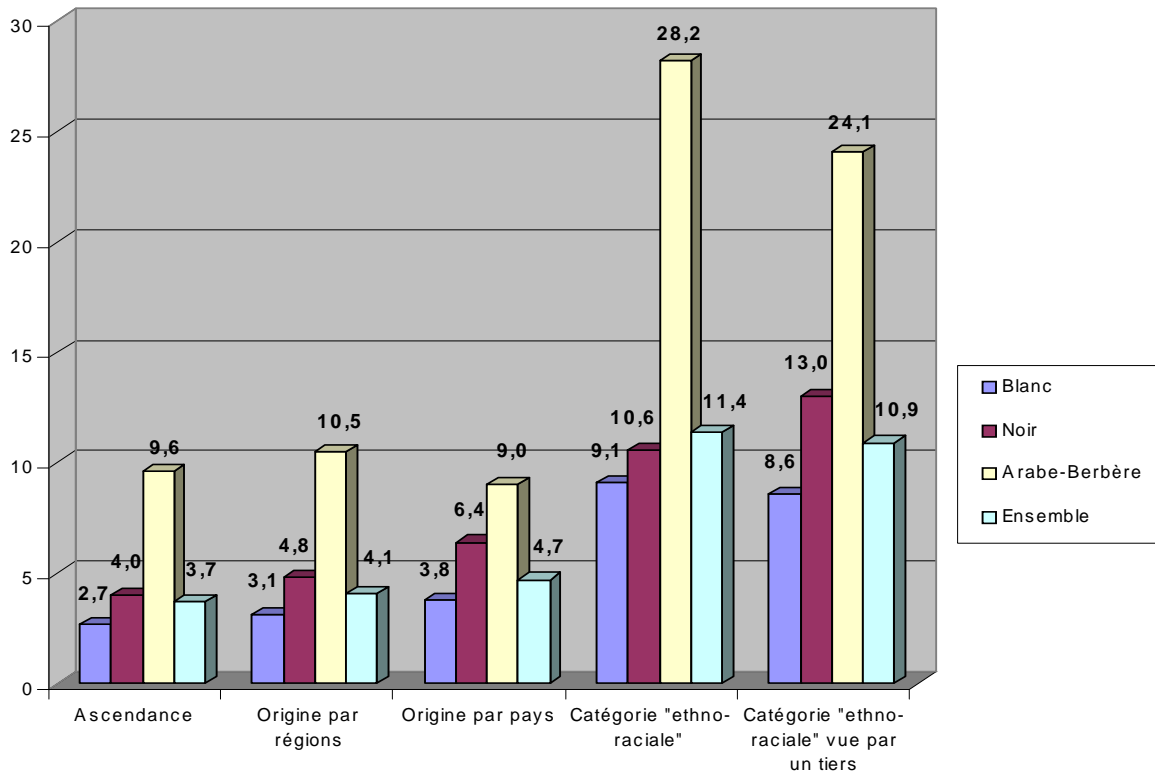
Le degré d'acceptation des catégories

Nous avons enfin demandé aux enquêtés s'ils étaient à l'aise ou non pour répondre au type de questions qui étaient posées dans l'enquête selon les diverses formes de classification. Le niveau d'acceptabilité se situe autour de 85% lorsqu'on est dans les trois premiers types de questionnement, et de 75% lorsqu'on s'intéresse aux catégories ethno raciales. Ce niveau est particulièrement élevé lorsqu'on le rapporte au contexte français actuel. En dépit de leur caractère inhabituel et des présentations caricaturales de la question dans le débat public, les catégories « ethno-raciales » sont globalement acceptées. Notons néanmoins que dans la présentation de l'enquête, nous avons replacé ces questions dans le cadre de la lutte contre les discriminations et de la promotion de la diversité.

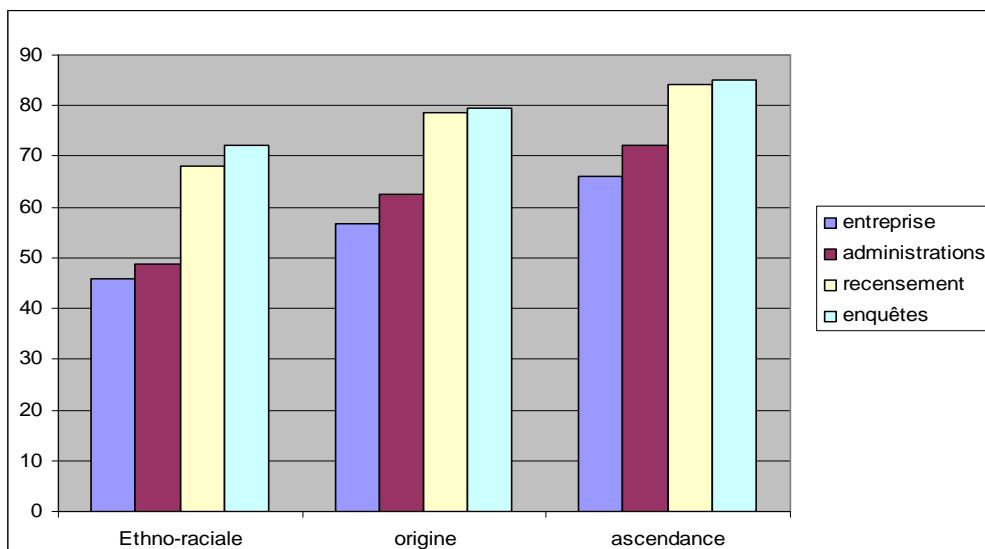
Réactions aux nomenclatures				
	Ascendance	Origine par régions	Origine par Pays	Catégories « ethno-raciales »
Sans problèmes	89	86.7	84.2	74.8
Mitigés	4.2	6.2	6.9	9.6
Réticents	3.8	3.9	4.6	12.4
NR	2.9	3.2	4.4	3.2
Total	100	100	100	100

Habituellement, les réponses à des questions de ce type varient beaucoup selon le sexe, l'âge, le niveau d'étude et la catégorie socio professionnelle. Dans notre cas, les variations sont faibles en fonction de ces caractéristiques. On observe en revanche des écarts significatifs en fonction de l'origine déclarée, et notamment pour les personnes s'étant classées comme « arabe ou berbère ». Le refus de ces questions est en effet plus marqué dans ce groupe que dans les autres. Mais même pour ce groupe de répondants, le niveau de refus explicite ne dépasse pas 30 %. Les personnes se déclarant « Noir » présentent un degré d'acceptation plus

élevé. On peut donc conclure à un effet –différencié– de l’origine sur l’acceptation des questions, mais l’enquête ne dit pas les raisons qui sous-tendent les refus ou les accords pour de telles questions².



Une forte variation existe également selon que ces catégories sont mises en œuvre au sein des entreprises ou des administrations – le refus est patent – ou au sein d’enquêtes ou du recensement – l’acceptation est très majoritaire.



² On notera également que les travaux méthodologiques sur l’acceptabilité de questions dans des enquêtes statistiques ne sont pas très développés. La comparaison avec le degré d’acceptabilité de questions sur les revenus par exemple montrerait qu’il est plus facile de donner son origine que de livrer son salaire.

Conclusion

Que retire-t-on de cette enquête ? Certains commentateurs ont voulu voir une acceptation relativement large des catégories notamment ethno raciales dans la société française ; à l'inverse, d'autres ont voulu y voir le signe d'un refus de ces mêmes catégories dans la société. Chacun doit se faire sa propre idée.

On peut toutefois indiquer qu'il paraît nécessaire aujourd'hui de construire des catégories permettant de dépasser les limites auxquelles nous allons très rapidement être confrontés avec les catégories actuelles :

1. C'est le cas des prénoms. Malgré la qualité des enquêtes qui ont été faites à partir des prénoms, on sait que ces derniers ont eux-mêmes une variabilité sociale extrêmement forte, et qu'ils ne sont pas représentatifs des populations qui sont exposées aux discriminations.

2. C'est le cas également du pays de naissance des parents. Cette donnée apparaît solide dans un temps encore limité : les personnes descendantes directement d'immigrés sont aujourd'hui en train de vieillir, elles ont entre 18 et 50 ans pour la plupart d'entre elles, elles ont déjà des enfants, et ces enfants-là ne sont plus des descendants directs d'immigrés, ce sont des petits-enfants d'immigrés. On risque de rencontrer de grandes difficultés pour décrire quels sont, parmi les petits enfants de l'immigration des années 60, ceux qui sont exposés aux discriminations. Seront-ils toujours exposés aux discriminations, et si c'est le cas, comment fera-t-on pour arriver à les mettre en évidence ?

Enfin, ce type de données inscrites dans des enquêtes spécifiques se différencie très nettement du monitoring d'entreprise. Il est évident que du côté des entreprises, dans le monde du travail, mais c'est valable aussi pour l'école et pour le logement, il convient de s'interroger sur la façon dont on pourrait obtenir des informations sur les traitements discriminatoires qui peuvent s'y produire. Si des enquêtes spécifiques peuvent permettre d'améliorer la connaissance des mécanismes responsables des discriminations, la surveillance (monitoring) des pratiques ordinaires passe par un système de veille statistique d'une autre nature. Il faut dès lors mettre en débat les différents protocoles afin de savoir jusqu'où il est possible d'aller ... sans aller trop loin.

Fiche signalétique

Intervenant : **Patrick Simon (Ined)**

Nom de l'opération : **Enquête « Mesure de la diversité »**

Source des données : Enquête spécifique, réalisée par l'INED en concertation avec un groupe de travail lié aux entreprises, financée par le FASILD à la demande du Ministère pour l'égalité des chances

Lieux et dates de la collecte : novembre 2005 – mars 2006 dans 7 entreprises (Axa, SNCF, Eau de Paris, L'Oréal, Adecco, Randstad, ED), 3 universités (Villetaneuse, Evry, et Paris 8), un IUFM et le Conseil Régional d'Ile-de-France.

Modes de recueil : L'enquête a utilisé un questionnaire auto-administré rempli devant l'équipe d'enquête. Pour les entreprises, sur le lieu de travail des salariés au cours de séances collectives (10 à 150 enquêtés), pour la SNCF, sur le lieu des tests d'embauche des candidats, pour les étudiants, à la fin d'un cours. Pour les entreprises de travail temporaire, les personnes ont été enquêtées par internet en accès sur un site sécurisé à l'INED.

Population ou périmètre concernés :

<i>Etablissements</i>	<i>Nombre de questionnaires</i>	<i>Taux de réponse</i>
L'OREAL	337	Personnes interrogées : Tout salarié des services concernés (exhaustivité). Un choix raisonné de services a été établi avec les entreprises
ADECCO (collecte par Internet)	169	
ED (GROUPE CARREFOUR)	68	
SNCF	219	
RANDSTAD (collecte par Internet)	155	Les taux de réponse sont variables. En dépit du volontariat mis en avant lors de la présentation de l'enquête, ils sont proches de 95% dans les entreprises et universités avec questionnaires auto-administrés, se situent autour de 50% pour les questionnaires sur internet. Seul le Conseil Régional d'IDF a connu un très faible taux de participation dû aux conditions de réalisation de l'enquête.
AXA	111	
EAU DE PARIS	71	
UNIVERSITE PARIS 8	42	
UNIVERSITE EVRY VAL ESSONNE	61	
UNIVERSITE DE VILLETANEUSE	24	
IUFM PARIS	33	
CONSEIL REGIONAL IDF	37	
TOTAL	1327	

Données utilisées pour construire les catégories d'"origines" :

L'enquête portait précisément sur les différentes catégorisations possibles de « l'origine ». Trois types d'enregistrement ont été testés : par ascendance (pays de naissance et nationalité des individus, de leurs parents et grands-parents), par auto-identification dans une liste de grandes régions (Maghreb, Antille-Caraïbe, Asie, Afrique, Europe, Amérique du Sud, Autre) ou de pays (France, Sénégal, Italie, Portugal, Pologne, Algérie, Turquie, ...) et enfin par auto-identification dans une nomenclature « ethno-raciale » (Blanc, Noir, Arabe ou berbère, Asiatique, Sous-continent Indien, Métis, Autre). Pour les deux dernières séries de questions, les réponses multiples étaient possibles (et c'était annoncé).

Références de publications

- Simon et Clément « Comment décrire la diversité des origines en France ? », *Population & Sociétés*, n°425, juil.-août 2006, www.ined.fr/fr/ressources_documentation/publications/
- Simon et Clément « Rapport 'Mesure de la diversité' », *Docs de travail de l'INED*, n°139, 2006, www.ined.fr/fr/ressources_documentation/publications/documents_travail/

Principaux résultats

L'enjeu central de l'enquête consistait à faire évaluer par les enquêtés les différentes nomenclatures proposées. D'une façon générale, les catégorisations par l'ascendance et l'origine géographique sont bien reçues par 86% des enquêtés qui n'éprouvent pas de difficulté à y répondre. Les susceptibilités sont plus marquées à l'égard des catégories « ethno- raciales », auxquelles près des ¾ des enquêtés n'éprouvent pas de difficulté à répondre. Le saut qualitatif entre les différentes approches est donc clairement marqué par les réactions des enquêtés. Cependant, si l'on considère le contexte français caractérisé par une faible familiarité avec les catégories « ethno- raciales », voire une forte hostilité à cette référence dans les débats publics, l'acceptation de cette nomenclature par près des ¾ des enquêtés montre une évolution très sensible des mentalités.

Le sexe, l'âge, le niveau de diplôme ou la position sociale ne produisent pas d'écarts très significatifs. La position dans la nomenclature « ethno- raciale » ou dans l'ascendance sur deux générations apportent des variations beaucoup plus sensibles. Les immigrés (16%) et leurs descendants directs (20%) ont une susceptibilité bien supérieure à la moyenne (12%) face aux catégories « ethno- raciales ». Une réticence similaire est enregistrée pour les personnes se considérant comme « Arabes ou berbères » (28%), tandis que les « Noirs » (11%) comme les « Blancs » (9%) ratifient plus facilement la nomenclature. On pourrait penser que, d'une façon générale, ce sont les minorités exposées aux discriminations qui se défient des catégorisations stéréotypées, mais la position singulière des personnes se considérant comme « Noires » oblige à relativiser le constat. Parmi les descendants d'immigrés ou les immigrés, ce sont les personnes d'origine maghrébine qui se montrent les plus réservées quant à la démarche. Cependant, aucun des sous-groupes identifiés dans l'enquête ne se prononce majoritairement contre l'une ou l'autre des grilles proposées.

Cette position générale plutôt favorable se module sensiblement lorsque des situations concrètes sont évoquées. Les fichiers d'entreprises et administratifs font l'objet d'une relative méfiance, tandis qu'une plus grande latitude est laissée aux enquêtes scientifiques et au recensement. Par exemple, alors que les questions sur l'ascendance et l'origine étaient unanimement acceptées, elles posent des problèmes à 20% des enquêtés si elles devaient figurer dans les fichiers du personnel des entreprises et l'enregistrement de catégories « ethno- raciales » susciterait l'hostilité du tiers des enquêtés. Là encore les clivages observés selon la position dans la nomenclature « ethno- raciale » s'approfondissent. Les personnes se considérant « Arabes ou berbères » expriment des réticences deux fois plus fréquentes que les « Noirs » ou les « Blancs » et sont à 51.5% réticents à l'introduction dans les fichiers du personnel de questions faisant référence aux catégories « ethno- raciales ».

Cette réserve exprimée à propos des fichiers d'entreprise se lève en partie lorsque la possibilité de définir des garanties sur l'utilisation des statistiques ainsi collectées ou sur l'intervention d'un organisme indépendant de l'entreprise est mentionnée. Il ne reste plus que 27% des enquêtés à déclarer ne jamais vouloir répondre à ce type de question dans le cas de la nomenclature « ethno- raciale » (37% des « Arabes ou berbères », 21% des « Noirs » et 25% des « Blancs »).

A contrario, la perspective de poser de telles questions dans les enquêtes scientifiques ou le recensement -résultat plus étonnant- suscite une large acceptation quelles que soient les nomenclatures envisagées. Même les catégories « ethno- raciales » sont reçues sans problèmes par 72% des enquêtés dans les enquêtes scientifiques et 68% encore dans le recensement. Les variations selon les caractéristiques socio- démographiques ou la position dans la nomenclature « ethno- raciale » restent relativement faibles, soulignant le consensus obtenu.

« Testing sollicité » au sein du Groupe Casino

Fabrice Foroni, ISM CORUM

Une vaste opération de « testing sollicité », conduite par ISM CORUM, est actuellement en cours [octobre 2007] au sein du Groupe Casino. Il y a donc certains points que je n'aborderai pas dans cette présentation :

- dans la mesure où l'opération de testing n'est pas terminée, je n'indiquerai pas certains détails qui pourraient risquer d'éventer les tests qui restent à réaliser ;
- un engagement a été pris, avec la direction et les organisations syndicales du Groupe Casino, de ne pas communiquer de résultats intermédiaires avant la fin de l'opération de « testing sollicité », prévue en juin 2008 : ce n'est qu'à ce moment-là que les résultats seront présentés, en premier lieu à la direction et aux syndicats du Groupe Casino, avant d'être diffusés plus largement.

[Nota : Les présents actes du colloque d'octobre 2007 ont été finalisés à l'automne 2008. Un post-scriptum a donc pu être adjoint à cette présentation, ci-après page 76 et suivantes, pour apporter des compléments de méthode et une synthèse des principaux résultats du testing, qui ont entre-temps été rendus publics, en juillet 2008]

Objets et finalités de l'opération

L'opération a pour objectif d'évaluer par testing les procédures de recrutement au sein du Groupe Casino, afin de vérifier si ces procédures donnent lieu à des discriminations en raison de « l'origine évoquée » par la consonance des nom et/ou prénom des candidats, ou dans certains cas par leur apparence (en envoyant un CV avec une photographie).

Chaque test consiste à faire postuler, sur un même emploi, deux profils de candidats comparables du point de vue de leur sexe (avec des paires de candidat.e.s tantôt masculines, tantôt féminines), de leurs compétences professionnelles et niveau d'expérience (en proposant des paires tantôt débutantes, tantôt confirmées), de leur niveau de formation, de leur lieu de résidence, de leur âge, de leur situation familiale et de leur nationalité (tous les candidats étant systématiquement de nationalité française, afin d'écartier toute différence liée au statut légal des candidats et d'invalider tout argument des recruteurs relatifs aux permis de séjour ou de travail).

Pour ne pas attirer l'attention des services en charge du recrutement, des différences mineures sont introduites entre les deux candidatures, tout en veillant à ce que ces différences ne puissent pas, sur le fond, faire pencher de façon légitime les choix des recruteurs. Des permutations systématiques sont par ailleurs effectuées d'un test à l'autre, afin d'alterner, entre les deux candidats de chaque paire, le contenu professionnel des CV ainsi que l'ordre dans lequel ils sont envoyés, de façon à contrôler statistiquement les risques de biais au niveau des résultats.

La seule différence qui puisse distinguer les deux candidats concerne « l'origine » que l'un et l'autre évoquent. Chaque test consiste à faire postuler un candidat dit « majoritaire », car évoquant par la consonance de ses nom et prénom une origine « hexagonale ancienne », tandis que l'autre candidat, dit « minoritaire », évoque une origine « extra-européenne ».

L'expression de « testing sollicité » souligne que c'est la direction du Groupe Casino, en accord avec les organisations syndicales, qui a elle-même souhaité une campagne de tests de discrimination au sein du Groupe, avec l'objectif d'évaluer les recrutements afférant à différentes filières professionnelles, sur des postes d'employé et d'encadrant. La direction et les organisations syndicales du Groupe Casino ont donc été parties prenantes dans la définition des objectifs et du périmètre des tests, dans le cadre d'un comité de suivi se réunissant de manière régulière. Cette démarche partenariale s'est construite dans le cadre du projet européen Equal nommé AVERROES (Actions Visant l'Egalité sans distinction de Race de Religion ou d'Origine dans l'Emploi et les Services).

En outre, les directions des ressources humaines des différentes sociétés du Groupe Casino ont été préalablement informées de l'opération de « testing sollicité » et de la possibilité que leurs services soient testés à n'importe quel moment. Ceci afin d'inciter les responsables à contrôler le fonctionnement non-discriminatoire des services dont ils ont la charge, et de sensibiliser les professionnels au respect de l'égalité de traitement dans leurs pratiques de travail. Il s'agit donc d'utiliser aussi la technique du testing comme un outil de management.

Pour autant, la réalisation des tests de discrimination est de la seule compétence d'un expert extérieur, ISM CORUM, qui a construit le protocole, qui met en œuvre la procédure concrète des tests, et qui est le seul à décider et à savoir où et quand les tests ont lieu.

La suite de la présentation adoptera le plan suivant :

1. Ce sont la direction et les organisations syndicales de Casino qui ont décidé de manière partagée le périmètre à tester (sociétés et types d'emploi). Je commencerai par préciser ce périmètre.
2. Je présenterai ensuite certains aspects de la mise en oeuvre des tests, et notamment les contraintes à prendre en compte pour faire en sorte que les tests passent inaperçus et s'apparentent à de vraies candidatures auprès des services et des personnes en charge de recruter.
3. Je conclurai ma présentation en précisant le format des résultats à attendre, et l'intérêt d'une telle opération de « testing sollicité » pour lutter contre les discriminations à l'embauche.

Périmètre de l'opération

La principale demande qui nous a été faite par la direction de Casino a été que le testing porte sur l'ensemble des sociétés du Groupe, afin que tout le monde se sente concerné par la question des discriminations.

Le périmètre du testing est donc composé d'une centaine d'hypermarchés *GEANT*, de plus de 300 supermarchés, 170 cafétérias, du réseau des *PETIT CASINO*, d'une trentaine d'entrepôts sous l'enseigne d'*EASYDIS*, des plates-formes de service-après-vente, ainsi que des services administratifs au siège du Groupe et de la centrale d'achats *EMC*. Cela représente au total près

de 700 sites, répartis dans environ 80 départements, même si les deux tiers des effectifs sont concentrés en Rhône-Alpes, PACA et Ile-de-France.

Une autre exigence a été que les tests portent sur l'éventail le plus large possible des types d'emplois présents dans le Groupe et sur des recrutements de tous niveaux, employés, agents de maîtrise et cadres.

D'où une grande diversité de métiers à tester, avec les métiers "classiques" de la grande distribution : employé libre service, caissier, magasinier, chef de rayon, manager... mais aussi des métiers de bouche (boulangier, pâtissier, boucher...), de vendeur spécialisé (hi-fi ou électro-ménager), de technicien-réparateur en SAV, de serveur ou de responsable de cafétéria, de préparateur de commande en entrepôt, de délégué commercial ou de gérant de *PETIT CASINO*...

Concernant la période de réalisation des tests, il a été décidé que ceux-ci s'étaleraient sur une année. Ce choix, rendu possible par le fait que l'étude s'inscrivait dans le cadre d'un projet européen Equal, a été motivé par plusieurs raisons :

- cette durée permet de prendre en compte tous les types de recrutement, y compris saisonniers ;
- il peut s'écouler plusieurs mois entre l'envoi des CV et le moment où les candidat.e.s seront sollicité.e.s ;
- une période d'une année permet en outre d'espacer les tests à destination d'un même service en charge du recrutement, et de faire en sorte qu'ils passent davantage inaperçus que s'il fallait les concentrer sur une période plus courte (point sur lequel je reviendrai plus loin).

Dernier aspect concernant le cadrage du testing : il a été décidé de limiter les tests à la phase initiale de sélection des CV, sans poursuivre jusqu'aux entretiens qui seraient éventuellement proposés aux candidats.

Cette restriction tient d'abord à des raisons pratiques, dans la mesure où la diversité des emplois à tester nécessiterait de disposer d'un éventail important de candidats testeurs qui soient crédibles. En outre, cela aurait impliqué des déplacements dans toute la France pour que les candidat.e.s se rendent aux entretiens, avec une très grande réactivité, notamment pour les candidat.e.s à un poste d'employé, qui doivent en principe résider à proximité de l'établissement où ils/elles postulent.

Dans un précédent testing réalisé en 2006 par ISM CORUM à la demande du Ministère de l'Emploi et du Bureau International du Travail³, les tests avaient concerné l'ensemble du processus de recrutement, y compris en envoyant des candidats aux entretiens d'embauche. Or, les résultats obtenus ont montré que la plus grande partie des discriminations de la part des employeurs s'opérait lors de la sélection des CV, avant même de rencontrer les candidat.e.s.

³ E. Cédiey et F. Foroni, *Les discriminations à raison de « l'origine » dans les embauches en France. Une enquête nationale par tests de discrimination selon la méthode du Bureau International du Travail*, Cahiers des Migrations Internationales n°85, OIT, 2007, 121 p. Les résultats de ce testing ont été rendus publics en mars 2007 et sont téléchargeables sur www.ismcorum.org

Le fait de restreindre ici les tests à la phase de sélection des CV ne devrait donc pas empêcher de repérer – s'il y a lieu – d'éventuelles discriminations. A condition de préparer avec soin les tests afin qu'ils se confondent avec de vraies candidatures et passent inaperçus...

Préparation et réalisation des tests

Il a d'abord fallu s'informer en amont des procédures de recrutement au sein du Groupe Casino, afin de préciser :

- les postes à pourvoir pour chacune des sociétés du Groupe et l'évolution saisonnière des besoins ;
- les modes de diffusion des offres : site internet de Casino, sites spécialisés, intermédiaires de l'emploi, presse quotidienne régionale, journaux gratuits... ;
- les modalités de candidature : spontanée ou en réponse à une offre, en remplissant un formulaire en ligne, par courrier postal, par téléphone, par présentation directe... ;
- la gestion des CV reçus : reroutage éventuel, personnes et services en charge de la sélection, critères de sélection, délais de réponse...

Comme chaque société du groupe Casino avait des procédures de recrutement spécifiques, une dizaine de réunions techniques ont dû être effectuées avec les responsables RH de chacune de ces sociétés. On peut néanmoins distinguer deux principaux types de recrutement :

- pour les postes d'employé, les candidatures sont le plus souvent spontanées (très peu d'offres d'emploi), en envoyant un CV directement aux établissements concernés (hypermarchés, supermarchés, cafétérias, entrepôts...), qui se chargent ensuite d'effectuer leur sélection ;
- pour les postes d'encadrant, le recrutement est centralisé au niveau de chaque société du Groupe, avec un ou plusieurs recruteurs en charge d'un territoire donné, et qui s'occupent d'examiner les CV reçus – le plus souvent par internet – en réponse à une offre ou par candidature spontanée.

Pour que les candidatures fictives du testing se confondent avec les vraies, des précautions ont également dû être prises lors de la mise en oeuvre des tests.

Les noms et prénoms utilisés dans les candidatures sont choisis en fonction de leur fréquence dans les générations auxquelles sont censé.e.s appartenir les candidat.e.s. En outre, il est apparu indispensable de disposer de plusieurs noms et prénoms utilisables pour évoquer les différentes origines testées (avec 160 noms et prénoms différents au total), de manière à ce qu'une même identité ne soit jamais utilisée deux fois pour tester un établissement ou service donné, ni même sur deux établissements géographiquement proches ou auprès de deux recruteurs de la même société, dans la mesure où il arrive qu'ils se transmettent des CV.

Pour les tests sur des postes d'employés, comme les candidats sont censés résider à proximité des établissements où ils postulent, il faut faire en sorte que les CV – qui ne peuvent être envoyés que par courrier postal – soient postés à proximité des établissements testés. Par ailleurs, pour recueillir les propositions éventuelles, il faut bien évidemment attribuer aux candidats de vrais numéros de téléphone, et pour les candidats à un poste de cadre, de vraies adresses emails.

Une autre mesure permettant de faire en sorte que les tests passent inaperçus consiste à tester de manière ponctuelle chaque service et personne en charge du recrutement durant les douze mois que dure l'opération. Par exemple, pour tester les recrutements sur des postes d'employés dans les 107 hypermarchés du Groupe Casino, chacun des 107 responsables RH concernés ne recevront au total que 4 tests en un an (1 test consistant à envoyer 2 CV).

Le tableau suivant précise les ratios fixés pour les sociétés et types d'emploi concernés par l'opération, ainsi que le nombre total de tests à réaliser (dernière colonne du tableau).

Type d'emploi et société	Nb de recruteurs	Nb de tests	... soit au total
<u>Employé en :</u> - Hypermarchés	107	x 4	= 428 tests
- Supermarchés - Cafétérias - Entrepôts - Serca (SAV)	552	x 2	= 1 104 tests
<u>Agent de maîtrise et cadre en :</u> Hyper, Super, Cafétéria, Proxi, Entrepôt, Siège...	une vingtaine de recruteurs	x 5	= 100 tests environ

(environ 1 600 tests au total)

Ces ratios permettent de faire en sorte que les quelques 1 600 tests à réaliser au total conservent un caractère ponctuel au niveau de chaque service ou personne en charge du recrutement, et viennent ainsi se fondre dans la masse des candidatures habituellement reçues. Pour indication :

- un hypermarché reçoit fréquemment plus d'une centaine de CV par mois ;
- un recruteur cadre a souvent à traiter une cinquantaine de CV en réponse à une offre d'emploi de manager.

Le caractère ponctuel des tests évite en outre de parasiter l'activité des services RH en les submergeant de faux CV.

La majorité des 1600 tests programmés portent sur les recrutements d'employés. En effet, ceux-ci représentent la plus grande partie des embauches réalisées chaque année par le Groupe Casino. D'après le Bilan Social 2006, les branches Hyper et Super ont ainsi effectué :

- 4800 embauches en CDI, dont 80 % sur des postes d'employés
- et environ 20 000 embauche en CDD, là aussi pour la plupart sur des emplois d'exécution.

Cette pondération est également liée aux craintes d'un faible taux de réponse pour les tests sur des postes d'employés, dans la mesure où il s'agit exclusivement de tests par candidatures spontanées. Tandis qu'un taux de réponse plus élevé est attendu pour les recrutements de cadres, lié au fait que même lorsque l'offre est déjà pourvue, les recruteurs s'intéressent en principe quand même aux CV qui continuent d'arriver, en raison des besoins quasi permanents sur les profils de manager dans le secteur de la grande distribution.

Format des résultats et intérêt de l'opération

Chaque test vise à placer un recruteur en situation de faire un choix entre deux candidat.e.s qui ne diffèrent que par leur « origine ». Le recueil des réponses à ces candidatures peut donner lieu à trois cas de figure.

1. Aucun.e des deux candidat.e.s n'est contacté.e. Le test est alors considéré comme non-exploitable, dans la mesure où il est probable que le recruteur n'a pas été placé en situation d'examiner les deux candidatures, soit parce que son choix s'est déjà porté sur un autre candidat bien réel, soit qu'il n'avait besoin de personne au moment où il a été testé (on a vu précédemment que la plupart des tests consistent à effectuer des candidatures spontanées).

2. La même proposition est faite aux 2 candidat.e.s : il s'agit ici du résultat en principe attendu, dans la mesure où les 2 CV étaient identiques. Pour les tests conduisant à un tel résultat, on considèrera que l'origine n'a apparemment pas influencé la sélection des CV. L'opération de testing sollicité n'a donc pas pour seul objectif de focaliser sur d'éventuels dysfonctionnements. Elle peut aussi permettre de mettre en valeur la qualité des recrutements en terme d'égalité de traitement et de souligner l'existence de bonnes pratiques.

3. Un.e candidat.e reçoit une proposition favorable, tandis que l'autre n'est pas retenu.e. Il ne s'agira pas dans ce cas de conclure à partir d'un seul test à l'existence d'une pratique discriminatoire. En revanche, si de tels choix ont tendance à se répéter sur plusieurs tests et à pencher en faveur d'un type de candidat.e plutôt qu'un autre, avec des écarts statistiquement significatifs, on pourra légitimement conclure à l'existence de phénomènes discriminatoires en raison de l'origine des candidat.e.s.

Ces résultats conduiront à faire un état des lieux des procédures de recrutement. Les taux de réponse obtenus à mi-parcours indiquent que l'on devrait disposer à l'issue de l'opération de plus de 400 tests exploitables – c'est-à-dire avec une réponse pour l'un.e ou pour les deux candidat.e.s. Ce qui permettra de procéder de manière assez fine à un état des lieux des pratiques de recrutement au sein du Groupe Casino, en mesurant l'intensité des écarts discriminatoires à l'embauche selon par exemple :

- la société concernée (hypermarchés, supermarchés, petits casinos, cafétérias, etc.) ;
- le type d'emploi (d'encadrement ou d'exécution, en rapport ou non avec la clientèle...)
- la taille, l'implantation (périphérie / centre-ville) et l'ancienneté du site testé ;
- l'expérience professionnelle affichée par les candidats (débutants ou confirmés) ;
- l'ordre d'envoi des deux CV.

Ces analyses devraient conduire à repérer de manière fine les bonnes pratiques éventuelles, ou au contraire des situations indiquant l'existence d'écarts de traitement en raison de l'origine des candidat.e.s.

Cette opération permettra en outre de disposer d'un outil de management :

- pendant toute la durée du testing sollicité, le fait que les services en charge du recrutement soient informés qu'ils peuvent être testés à tout moment permet de les sensibiliser au respect de l'égalité de traitement dans leurs pratiques de recrutement ;
- une fois l'opération terminée, il sera possible d'approfondir ce travail de veille, et si besoin d'amélioration des pratiques de recrutement, à partir des résultats obtenus et grâce à une traçabilité complète de tous les tests effectués ;

- enfin, l'opération sera susceptible d'être reconduite par la suite à l'identique, ou de manière plus restreinte, en resserrant par exemple sur des secteurs où des dysfonctionnements seraient apparus, et les résultats des tests pourront alors s'inscrire dans une logique de mise en œuvre d'indicateurs de suivi, de tableaux de bord et de « contrôle qualité » des procédures de recrutement.

Fiche signalétique

Intervenant : **Fabrice Foroni (ISM-Corum)**

Nom de l'opération : « **Testing sollicité** » au sein du **Groupe Casino**

<i>Source des données :</i>	Enquête spécifique
<i>Lieu(x) de la collecte :</i>	France entière
<i>Mode(s) de recueil :</i>	Enregistrement des réponses (par téléphone, email ou courrier postal) obtenues après l'envoi (par paire) des CV de candidats fictifs
<i>Périmètre concerné :</i>	Toutes les sociétés, services et personnes en charge du recrutement sur des postes d'employé ou d'encadrement au sein du Groupe Casino (hyper et supermarchés, petits casinos, cafétérias, entrepôts, services après-vente, centrale d'achat, siège)
<i>Taille de l'échantillon :</i>	Environ 1600 tests (un test = envoi d'une paire de CV) engagés entre mars 2007 et mars 2008
<i>Taux de réponse :</i>	34 % de tests « exploitables » (= proposition faite à l'un et/ou à l'autre candidat.e.s) pour les 700 tests engagés avant le 01/08/2007
<i>Donnée(s) utilisée(s) pour construire les catégories d'"origines" :</i>	Les candidat.e.s, tou.te.s de nationalité française, évoquent des « origines » non hexagonales par la consonance de leurs noms et/ou prénoms ou parfois par une photo.
<i>Libellés de ces catégories :</i>	Informations non communicables en octobre 2007 car des tests sont en cours.
<i>Principales questions explorées en lien avec les variables d'"origines" :</i>	Calcul de taux de sélection suivant « l'origine évoquée » par les candidat.e.s, afin de repérer d'éventuels écarts de traitement ou bien l'absence d'écart. Analyse des variations de ces taux suivant notamment : <ul style="list-style-type: none">- la société concernée (hypermarchés, supermarchés, petits casinos, cafétérias...)- le type d'emploi postulé (d'encadrement ou d'exécution, en lien ou non avec la clientèle...)- la taille, l'implantation (périphérie / centre-ville) et l'ancienneté du site où l'emploi est à pourvoir- l'expérience professionnelle affichée par les candidat.e.s (débutant.e.s ou confirmé.e.s)
<i>Principaux résultats :</i>	non communicables avant juin 2008

Post-scriptum

Le testing sollicité au sein du Groupe Casino s'est achevé selon l'échéancier prévu. Ses résultats ont été rendus publics de façon complète début juillet 2008.

Compléments méthodologiques

1. La longue et minutieuse phase de préparation du contenu et du protocole des tests a apparemment été déterminante pour faire en sorte que ceux-ci, bien qu'annoncés au sein du Groupe Casino, se fondent parmi les candidatures réelles et passent inaperçus auprès des services testés. Aucun retour sur un test éventé n'est remonté des réseaux des différentes sociétés du Groupe. Les résultats des tests eux-mêmes accréditent ce constat.

2. Pour faire la paire avec le candidat évoquant une origine « hexagonale ancienne », cinq origines « extra-européennes » différentes ont servi aux tests – dans les mêmes proportions – auprès des services en charge du recrutement de toutes les sociétés. D'un test à l'autre, le candidat évoquant une origine « hexagonale ancienne » a ainsi postulé en même temps qu'un candidat qui pouvait évoquer :

- soit une origine « noire africaine » par la consonance de ses nom et prénom ;
- soit une origine « asiatique » par la consonance de ses nom et prénom ;
- soit une origine « maghrébine » par la consonance de ses nom et prénom ;
- soit une origine que nous appellerons « franco-maghrébine », par l'association d'un nom à consonance « maghrébine » et d'un prénom à consonance « hexagonale ancienne » ou d'usage mixte ;
- soit une origine que nous appellerons « antillaise », en joignant au CV la photographie d'une personne noire de peau et portant un nom et un prénom à consonance « hexagonale ancienne » – tandis que les CV de tous les autres candidats étaient sans photographie.

Synthèse des principaux résultats

Près de 1 600 tests – ce qui représente donc 3 200 candidatures – ont été effectués sur une période d'une année auprès de l'ensemble des supermarchés, hypermarchés, cafétérias, entrepôts et sites de service après vente du Groupe Casino implantés dans l'hexagone, afin de tester les recrutements effectués par chacun de ces établissements sur les postes d'employés. Des propositions ont été faites par les établissements testés en réponse à l'une et/ou l'autre candidatures dans 428 tests.

La répartition de ces 428 tests dans le tableau 1 montre que seulement 31 % des établissements testés ont contacté les deux candidats fictifs qui leur étaient proposés. Tandis que 50 % ont appelé uniquement le candidat dit « majoritaire », qui évoquait une origine « hexagonale ancienne », et que les 19 % restants n'ont contacté que le candidat dit « minoritaire », qui évoquait l'une des origines « extra-européennes ».

Cela signifie que les établissements qui n'ont appelé qu'un seul candidat ont choisi près de trois fois sur quatre le candidat évoquant une origine « hexagonale ancienne ». Au total, celui-ci a donc été contacté dans $31 + 50 = 81$ % des 428 tests – seul ou avec son binôme. Tandis

que le candidat évoquant une origine « extra-européenne » n'a été appelé que dans 31 + 19 = 50 % des tests, alors qu'il affichait pourtant des compétences et niveaux d'expérience comparables au candidat « majoritaire » postulant sur le même emploi.

Tableau 1 : Résultats d'ensemble des tests sur les recrutements d'employés

		(a)	(b)	(c)	(a + b)	(a + c)
<i>Total tests exploitables</i>		les deux contactés	"majoritaire" seul contacté	"minoritaire" seul contacté	"majoritaire" contacté	"minoritaire" contacté
(428)	100 %	31 %	50 %	19 %	81 %	50 %

Les tests effectués auprès des différents établissements du Groupe Casino indiquent donc d'importantes différences de traitement au détriment des candidats évoquant une origine « extra-européenne » sur les recrutements d'employés.

Dans le rapport complet, les résultats sont détaillés pour chacun des types d'emploi testés, permettant de constater que des écarts discriminatoires statistiquement significatifs s'observent aussi bien dans les recrutements d'employé commercial en supermarché ou en hypermarché, que sur les postes d'employé de restauration en cafétéria, de cariste ou de préparateur de commandes en entrepôt. En revanche, aucun écart de traitement statistiquement significatif n'a été constaté pour les recrutements effectués par les hypermarchés sur les postes de vendeur spécialisé en électro-ménager, hi-fi et informatique.

De même, les analyses détaillées dans le rapport montrent qu'aucun écart statistiquement significatif n'a été constaté dans les recrutements sur des postes d'encadrement – les tests ayant porté sur des offres d'emploi de manager commercial, chef de rayon, directeur adjoint de cafétéria, contrôleur de gestion...

Il est apparu par ailleurs que les écarts discriminatoires les plus importants ont été observés dans les tests posant deux candidatures fictives de niveau débutant.

Les résultats du tableau 2 montrent en effet que le candidat « minoritaire » a été contacté quasiment deux fois moins souvent que le candidat « majoritaire » lorsqu'ils affichaient tous deux un niveau débutant. Tandis que les stéréotypes à l'encontre des candidats évoquant une origine « extra-européenne » ont perdu en vigueur lorsque les deux candidats affichaient un niveau confirmé – les écarts se réduisant, tout en restant statistiquement significatifs.

Tableau 2 : Résultats suivant le niveau d'expérience affiché par les deux candidatures fictives

		(a)	(b)	(c)	(a + b)	(a + c)	
<i>Niveau d'expérience</i>		<i>Total tests exploitables</i>	les deux contactés	"majoritaire" seul contacté	"minoritaire" seul contacté	"majoritaire" contacté	"minoritaire" contacté
Paires débutantes	(157) 100 %	24 %	58 %	18 %	82 %	42 %	
Paires confirmées	(271) 100 %	35 %	46 %	19 %	81 %	54 %	

Paires "débutantes" = les deux candidat.e.s affichent une expérience de 1 à 3 années sur un poste similaire
Paires "confirmées" = les deux candidat.e.s affichent une expérience de 5 à 10 années sur un poste similaire

La variable du genre intégrée dans le profil des paires testées a permis de constater que les écarts discriminatoires en fonction de l'origine évoquée étaient de même ampleur entre deux candidats qu'entre deux candidates.

Des écarts importants ont également été constatés quelle que soit le type d'origine « extra-européenne » évoquée par le candidat « minoritaire », y compris pour des origines – telle qu'« asiatique » – qui ne sont pas a priori perçues dans les enquêtes d'opinion comme potentiellement victimes de discriminations.

Enfin, il est apparu que les résultats variaient également en fonction des caractéristiques des établissements testés, et notamment de l'ancienneté de leur population salariée.

Les résultats du tableau 3 montrent en effet que les écarts discriminatoires sont nettement plus importants pour les tests effectués auprès d'établissements ayant une population salariée avec une forte ancienneté moyenne. Tandis que ces écarts ont tendance à se réduire – tout en restant statistiquement significatifs – lorsque l'ancienneté moyenne de l'établissement testé est plus faible.

Tableau 3 : Résultats suivant l'ancienneté de la population salariée des établissements testés

Ancienneté de la population salariée	Total tests exploitables	(a) (b) (c)			(a + b) (a + c)	
		les deux contactés	"majoritaire" seul contacté	"minoritaire" seul contacté	"majoritaire" contacté	"minoritaire" contacté
forte	(156) 100 %	29 %	56 %	15 %	85 %	44 %
faible	(239) 100 %	32 %	47 %	21 %	79 %	53 %

N.B. : les supermarchés, hypermarchés, cafétérias, entrepôts ont été répartis en deux moitiés : l'une composée des établissements ayant l'ancienneté moyenne la plus faible ; l'autre étant composée des établissements ayant la plus forte ancienneté. 33 tests ne sont pas comptés ci-dessus car concernant des établissements affiliés ou franchisés, dont l'ancienneté moyenne est indéterminée.

D'autres comparaisons ont été effectuées en fonction de la commune d'implantation des établissements, ainsi que de leur localisation géographique sur le territoire hexagonal. Elles ont permis de constater que, contrairement à certaines idées reçues, les pratiques discriminatoires ne sont ni plus ni moins fréquentes en milieu rural qu'en milieu urbain, et qu'elles ne sont pas non plus l'apanage d'une région en particulier. Aucune tendance ni clivage d'ordre socio-géographique n'a ainsi pu être mis au jour.

Le rapport complet du testing sollicité par le Groupe Casino est téléchargeable sur le site d'ISM CORUM : <http://www.ismcorum.org/testing-sollicite-par-casino-M22-R22-A367.html>

L'expérience de l'ethnic monitoring aux Pays-Bas. Recensement interne des origines des salariés par le lieu de naissance de leurs parents et objectif de rééquilibrage de leurs proportions dans le cadre de la loi Samen, 1998-2003

Virginie Guiraudon, CNRS

Mon exposé porte sur une expérience assez particulière de « monitoring » des minorités ethniques, aux Pays-Bas. Les Pays-Bas et le Royaume-Uni ont été parmi les pays qui ont le plus développé les politiques publiques de lutte contre les discriminations, et qui ont inspiré en partie le texte de la directive européenne de 2000 visant à lutter contre les discriminations fondées sur l'origine ethnique et raciale.

1979, politique des minorités

En 1979, le WRR a produit un rapport montrant qu'il convenait de prendre en compte aux Pays-Bas la situation des minorités issues de l'immigration de travail et de l'immigration post-coloniale. Le WRR est le « Conseil scientifique pour la politique gouvernementale » (sorte de Commissariat général au Plan en plus indépendant et plus influent) qui est à l'origine d'un certain nombre de propositions auxquelles le gouvernement est légalement tenu de répondre. Suite à ce rapport, les Pays-Bas ont adopté, en 1983, une « politique des minorités » comprenant plusieurs volets.

Le volet que l'on évoque le plus souvent sous le nom de « le multiculturalisme néerlandais » visait la préservation des identités des différents groupes immigrés ou originaires des colonies néerlandaises. Mais cette politique faisait également référence à des notions telles que la promotion de l'égalité des chances, la lutte contre les discriminations non légales, l'égalité juridique. Un travail a d'ailleurs été conduit par des juristes qui ont nettoyé le droit néerlandais de toutes formes de discrimination directe et indirecte, entre étrangers, immigrés et néerlandais : 1200 lois, décrets, textes, ont ainsi été considérés comme discriminatoires envers les immigrés.

1994, l'intégration socio-économique des minorités ethniques

En 1990, le WRR revient sur ses positions concernant le multiculturalisme. C'est alors moins l'identité culturelle qu'il convient de promouvoir que l'intégration socio économique des minorités ethniques. Le WRR estime en effet que le vrai problème tient à la ségrégation spatiale, au chômage de ces populations, et au fait qu'elles n'accèdent pas à des niveaux d'étude suffisants et notamment aux filières prestigieuses. C'est donc l'intégration, terminologie empruntée au vocabulaire français, qui est mise en exergue. En 1994, Le gouvernement reprendra à son compte ce terme en privilégiant l'intégration dans et par la travail, par exemple en insistant sur l'apprentissage de la langue néerlandaise pour faciliter son insertion sur le marché de l'emploi.

La politique néerlandaise envers les « allochtones » s'appuie sur des groupes cibles. Certaines populations sont définies en fonction de leur nationalité. Pour d'autres, les catégories sont plus fines : par exemple, un indonésien né du temps des colonies sera affecté à un groupe, et un indonésien né quelques années plus tard après la décolonisation sera affecté à une autre catégorie. Enfin, certaines catégories de populations sont constituées à partir de l'origine : les personnes définies comme étant originaires du Surinam, des Antilles néerlandaises, de Turquie, du Maroc, sont des personnes nées dans ces pays ou ayant au moins un parent né dans ces pays.

Par ailleurs, les catégories des minorités tiennent également compte de la situation économique : ainsi, les allochtones non occidentaux excluent les japonais considérés comme n'ayant pas de problèmes particuliers. En somme, ce qu'on définit comme « minorité ethnique », c'est principalement une origine fondée sur le pays de naissance des parents, et ensuite une situation socio-économique considérée comme problématique.

Population par origine ethnique, 1997 (en milliers)		
NEERLANDAIS « AUTOCHTONES »		12.890
MINORITES ETHNIQUES	Turcs	280
	Marocains	233
	Surinamiens	287
	Antillais/Arubains	95
	<i>Grecs</i>	11
	<i>Italiens</i>	32
	<i>(Ex) yougoslaves</i>	60
	<i>Capverdiens</i>	17
	<i>Portugais</i>	13
	<i>Espagnols</i>	29
	<i>Tunisiens</i>	6
AUTRES "ALLOCHTONES"		1.613
TOTAL		15.567

En 1997, sur les 15 millions et demi de néerlandais, les minorités ethniques représentent un peu plus d'un million de personnes, sachant qu'il y a 1,6 millions d'autres allochtones considérés comme des groupes ne présentant pas de problèmes spécifiques. Sont en gras dans le tableau les groupes qui, dans les années 1980, étaient considérés comme ayant des problèmes.

Note : Le système statistique néerlandais

Il n'y a plus, aux Pays-Bas, de recensement depuis 1971 suite à une grande campagne conduite à la fin des années 1970 contre un recensement jugé contraire au respect de la vie privée. Depuis, le service de statistiques (le CBS) s'appuie sur l'exploitation des registres municipaux informatisés qui, paradoxalement, livrent un très grand nombre d'informations sur les néerlandais.

Sur les questions d'emploi, il existe également des enquêtes spécifiques particulièrement approfondies : ces enquêtes servent à informer ensuite les comparaisons européennes et Eurostat. On y trouve 5 catégories : les Néerlandais nés aux Pays-Bas de parents nés aux Pays Bas, et les quatre plus grands groupes numériques (Turcs, Marocains, Surinamiens, Antillais/Arubains) qui sont considérés comme ayant des problèmes particuliers d'intégration socio économique. Ces enquêtes sur l'emploi portent sur un échantillon très large ; faites par téléphone, elles n'apparaissent cependant pas toujours très fiables du fait que les enquêteurs ne parlent pas les langues d'origine. Les taux de chômage calculés à partir d'autres enquêtes effectués dans le cadre des politiques d'intégration et n'interrogeant que des personnes issues des minorités ethniques sont par exemple beaucoup plus importants. Ces enquêtes complémentaires sont conduites par des organismes para publics de recherche, en liaison avec la commande politique et la recherche universitaire.

1997, la loi SAMEN

Que se passe-t-il en 1997 ? La situation économique est très bonne, le taux de chômage des Néerlandais de parents néerlandais est de 4% : on parle à ce propos de miracle néerlandais. En revanche, le taux de chômage des allochtones (minorités ethniques et autres allochtones) est de 16%, et celui des minorités ethniques de 20%.

Face à cette situation, la loi SAMEN ou « Loi de promotion de l'emploi des minorités » est adoptée. « Samen » signifie également « ensemble » en néerlandais. Cette loi pose le cadre problématique général, elle pointe les grandes questions qu'il semble nécessaire de se poser à propos de l'emploi des minorités : est-ce un problème de demande ou d'offre, est-ce que ce sont les employeurs qui ont un problème, le système éducatif, que peut-on y faire ?

Cette loi met également en place un dispositif spécifique visant l'emploi des minorités ethniques. La loi s'applique aux entreprises de plus de trente cinq salariés – les entreprises ne souhaitant pas participer à cette loi et le motivant par écrit pouvaient toutefois y échapper. Le gouvernement publiait des proportions de minorités ethniques à atteindre par région. Chaque entreprise devait tendre, dans sa composition, à refléter les proportions des différentes minorités déclarées dans la région où l'entreprise était établie. L'employeur devait donc recueillir le pays de naissance de l'employé et de ses parents, afin de déterminer s'il faisait partie d'un groupe cible. Les qualifications étaient évidemment prises en compte.

Ce dispositif a été considéré par les autorités gouvernementales comme un succès : 70% des entreprises concernées se sont mises en conformité avec la loi. Ce qui est d'autant plus notable qu'un dispositif expérimental analogue n'avait séduit, deux ans auparavant, que 10% des entreprises.

L'action autour de la loi SAMEN

Chaque entreprise répondant à la loi SAMEN devait établir un rapport annuel indiquant la représentation des minorités ethniques en son sein (définies par le lieu de naissance des personnes et de leurs parents), l'évolution en cours d'année, et les mesures prévues pour améliorer cette représentation. Ce rapport était discuté par le « Conseil du personnel » de

l'entreprise, qui émettait un avis. Le rapport était ensuite publié et déposé à l'agence pour l'emploi néerlandaise.

Avant la loi SAMEN, les « Conseils du personnel » ne s'intéressaient que très peu à ces questions, privilégiant les problématiques relatives aux conditions de travail ou à l'égalité professionnelle entre les hommes et les femmes. Ils se sont rapidement investis dans ce nouveau rôle. Les organisations syndicales ont également pu investir un rôle d'alerte.

De même, les résultats des entreprises étant publiés, la société civile a été en mesure d'exercer des pressions ciblées sur les entreprises s'écartant de leurs objectifs. Certaines associations se sont ainsi appuyé sur les statistiques du personnel des entreprises pour pointer les « bonnes pratiques » ou pour rencontrer les entreprises en défaut – moins dans l'idée de les condamner que, suivant une démarche de médiation, dans l'idée de les accompagner et d'analyser avec elles les problèmes qu'elles pouvaient rencontrer. Par exemple, l'ONG « Mira Media », qui travaille sur les minorités visibles dans les médias, a pu souligner que dans la télévision publique néerlandaise, l'objectif de minorités ethniques n'était pas atteint, contrairement aux acteurs privés comme MTV ou Endemol. De même « Radar », qui est une association importante de Rotterdam travaillant sur les discriminations, a développé une logique de médiation avec les entreprises en se basant sur ces données.

Enfin, la « Commission de l'égalité des chances », l'équivalent de la HALDE, s'est aussi servie de ces données dans l'idée de mesurer les discriminations indirectes, pour avoir une vision d'ensemble de la situation. La loi SAMEN a ainsi permis à un certain nombre d'acteurs de se saisir de ces questions.

Les résultats de la loi SAMEN

Le taux de chômage des minorités ethniques baisse à partir de 1997. Chacun a ainsi pu se réjouir des effets prêtés au dispositif de la loi SAMEN. Toutefois, une forte croissance était observable à cette époque, conduisant à l'embauche massive de chômeurs de longue durée et de jeunes issus des quartiers sensibles. Aussi est-il difficile de pondérer les effets de la loi et ceux de la conjoncture économique dans la baisse du chômage des minorités ethniques à ce moment-là.

D'un point de vue statistique, l'objectif global de 10% de minorités ethniques dans l'emploi n'a pas été atteint, notamment parmi les rémunérations élevées. Cependant il y a eu progrès : toutes minorités, toutes catégories professionnelles, tous secteurs et toutes régions confondus, le résultat obtenu a été de 8,3%.

Le devenir de la Loi SAMEN

Reconduite une première fois, la loi a été finalement abandonnée en 2003, un certain nombre d'acteurs estimant que les résultats étaient suffisants et qu'il n'y avait plus de nécessité de poursuivre dans cette voie.

Les associations d'employeurs figuraient parmi les acteurs poussant à l'abandon de la loi SAMEN. Elles ont trouvé un fort relais parmi les acteurs politiques. En effet, plusieurs partis de la coalition au gouvernement ont, dès la fin des années 1990, développé un discours

empreint d'un fort accent xénophobe et islamophobe. Les minorités ethniques ont dès lors été construites comme posant problème et les actions positives en leur faveur ont été délégitimées.

Par ailleurs, certains salariés appartenant aux publics cibles n'ont pas apprécié d'être stigmatisés à travers leur labellisation comme « minorité ethnique ». Le contexte politique extrêmement xénophobe conduisant à prêter aux minorités des caractéristiques négatives, nombreux ont été ceux souhaitant échapper à cette catégorisation. C'était le cas, en particulier, des indonésiens, qui se considéraient non pas comme une minorité ethnique mais comme un groupe rapatrié au moment de la décolonisation.

De l'autre côté, les syndicats, les associations anti-racistes, les Verts, le Parti social-démocrate et la Commission de l'égalité des chances souhaitaient que la loi soit prolongée. Ils attendaient notamment de pouvoir poursuivre l'effort engagé après que la loi SAMEN ait pu apporter un certain nombre de réponses en termes de diagnostics et d'identification des pratiques discriminatoires dans l'embauche et le déroulement de carrière.

La loi SAMEN a été supprimée au profit de ce qu'on appelle le « diversity management », le management de la diversité. Les entreprises souhaitant continuer à enregistrer l'origine de leurs employés peuvent le faire – en respectant un certain nombre de précautions déontologiques. Un « Centre du management de la diversité » géré par les employeurs et disposant d'un certain nombre d'ambassadeurs a été mis en place, mais l'obligation légale a dorénavant été levée.

Remarques finales

Sophie Vulliet-Tavernier, CNIL
directrice des affaires juridiques et de l'expertise

Je voudrais faire deux observations suite aux différentes interventions de cette journée.

Sur la méthode patronymique et l'analyse des prénoms, la CNIL a fortement rappelé qu'en aucun cas le nom ou le prénom ne peut conduire à déterminer une origine ethnique ou raciale. Que l'analyse des noms et des prénoms puisse servir pour détecter d'éventuelles pratiques discriminatoires, dans le cadre d'une analyse multicritères, c'est une évidence. Mais pour le reste faisons attention à tout usage abusif ou stigmatisant, dans un contexte qui est très politisé et délicat et sur des enjeux qui sont extrêmement importants.

A ce titre je souhaiterais revenir sur la position de la CNIL en regard de l'enquête TIES diligentée par l'INED. Dans cette enquête, l'analyse des patronymes permettait de déterminer les descendants des migrants d'origine turque, liant donc nom et origine. Cette méthode a été admise, après d'âpres débats, par la CNIL, d'une part parce qu'aucune méthode alternative n'était envisageable, et d'autre part parce qu'un certain nombre de garanties concernant les droits des enquêtés ont été respectées (envoi d'une lettre d'information aux personnes sélectionnées, recueil de leur accord). Ces méthodes nécessitent donc un grand nombre de précautions.

Je voudrais aussi revenir sur les méthodes de testing. La CNIL n'est a priori pas concernée par cette méthode. Nous avons toutefois été saisis, il y a quelques mois, de plaintes concernant certaines modalités de recrutement de testeurs. Nous avons été amenés à engager des procédures de sanction à l'encontre de deux sociétés d'étude qui avaient créé des fichiers de recrutement de testeurs incluant des données sensibles rapportées par des tiers et portant sur l'orientation sexuelle et la couleur de peau des candidats. Cette collecte de données, à l'insu des intéressés, est tout à fait condamnable. Elle engage à être extrêmement vigilant sur les sociétés réalisant de telles études par testing.

En conclusion, la CNIL que je représente ici se félicite de la tenue de ce type de colloque. Vous avez su éviter toute politisation des débats, maintenir une excellente qualité scientifique, et je crois qu'à l'heure où l'on assiste à une politisation excessive sur cette question des statistiques ethniques, il est tout à fait bon d'avoir ce type d'échanges.

Principales références en ligne

Eric Cédiey, « Intervention au séminaire de l'unité Migrations Internationales et Minorités de l'INED : Testing, avancées et limites d'une méthode », 5 juin 2007 – ISM-CORUM : <http://ismcorum.pemacoh.com/UserFiles/File/intervention%20colloque%20Eric.pdf>

Eric Cédiey et Fabrice Foroni, *Un diagnostic partagé sur les discriminations liées à l'origine et au sexe. Résultat d'une recherche-action au sein du Groupe Casino*, 2005, ISM-CORUM : http://www.ismcorum.org/stock_images/reference/rapport-casino-ism-corum.pdf

Eric Cédiey et Fabrice Foroni, *Les discriminations à raison de l'origine dans les embauches en France. Une enquête nationale par tests de discrimination selon la méthode du Bureau International du Travail*, 2007, financement DARES

BIT : www.ilo.org/public/french/bureau/inf/download/discrim_france.pdf

ISM-CORUM : <http://www.ismcorum.org/testing-bit-france-2007-M22-R22-A361.html>

Eric Cédiey, Fabrice Foroni et Hélène Garner, 2008, « Discriminations à l'embauche fondées sur l'origine à l'encontre de jeunes Français(es) peu qualifié(e)s », *Premières Synthèses*, n° 06.3, DARES : www.travail-solidarite.gouv.fr/IMG/pdf/2008.02-06.3.pdf

Centre d'analyse stratégique, « Statistiques "ethniques" : éléments de cadrage », *Rapports et Documents*, n° 3, octobre 2006 : www.strategie.gouv.fr/article.php3?id_article=339

Centre d'analyse stratégique, « La discrimination saisie sur le vif : le testing », *Horizons stratégiques*, n° 5, juillet 2007 – www.strategie.gouv.fr/revue/revue.php3?id_rubrique=215

CNIL, Anne Debet (Rapporteur), « Mesure de la diversité et protection des données personnelles », Rapport présenté en séance plénière le 15 mai 2007 :

<http://www.cnil.fr/fileadmin/documents/approfondir/dossier/diversite/RapportdiversiteVD.pdf>

Commission européenne, *Etude comparative de la collecte de données visant à mesurer l'étendue et l'impact de la discrimination aux Etats-Unis, Canada, Australie, Royaume-Uni et Pays-Bas*, 2004, Office des publications officielles des Communautés européennes :

http://ec.europa.eu/employment_social/publications/2005/ke6204874_fr.pdf

Commission européenne, *La lutte contre la discrimination et la promotion de l'égalité. Comment mesurer les avancées réalisées*, 2008, Office des publications officielles :

http://ec.europa.eu/employment_social/fundamental_rights/pdf/pubst/stud/measprog08_fr.pdf

Marie-Noëlle Decharne et Eric Liedts, *Porter un prénom arabe ou musulman est-il discriminant dans l'enseignement supérieur ? Orientation et poursuite d'études dans l'enseignement supérieur du Nord - Pas de Calais*, mai 2007, ORES, Pôle Universitaire Lille Nord - Pas de Calais : http://www.poleuniv-lille-npdc.fr/n_telecharg/ores/04-05_B_R001.pdf

Emmanuel Duguet, Noam Leandri, Yannick L'Horty, Pascale Petit, « Discriminations à l'embauche. Un testing sur les jeunes des banlieues d'Île-de-France », *Rapports et documents*, 2007, Centre d'analyse stratégique : <http://www.strategie.gouv.fr/IMG/pdf/2007-03-05discriminationsembauche-testing.pdf>

Fabrice Foroni, *Résultats du testing sollicité par le Groupe Casino*, Rapport complet et Synthèse, 2008, ISM CORUM : <http://www.ismcorum.org/testing-sollicite-par-casino-M22-R22-A367.html>

Denis Fougère et Julien Pouget, « L'emploi public s'est-il diversifié ? Sexe, niveau d'étude, origine sociale et origine nationale des salariés de la fonction publique », Complément au rapport sur *La diversité de la fonction publique*, Dominique Versini, La Documentation Française : <http://lesrapports.ladocumentationfrancaise.fr/BRP/054000049/0000.pdf>

Alain Frickey, Jake Murdoch et Jean-Luc Primon, « Les jeunes issus de l'immigration. De l'enseignement supérieur au marché du travail », *Bref*, n° 205, février 2004 – CEREQ : <http://www.cereq.fr/pdf/b205.pdf>

NOVETHIC, Anne-Catherine Husson, « Le Groupe Casino établit le premier diagnostic partagé sur la discrimination », 29 juin 2005 : <http://www.novethic.fr/novethic/site/article/index.jsp?id=91859>

NOVETHIC, Véronique Smée, « Testing : Casino joue la transparence », 16 juillet 2008, <http://www.novethic.fr/novethic/v3/article.jsp?id=117076>

Julien Pouget, « La Fonction publique : vers plus de diversité ? », in *France, Portrait Social*, édition 2005-2006, INSEE : http://www.insee.fr/fr/ffc/docs_ffc/FPORSOC05c.PDF

Julien Pouget et Romain Aeberhardt, « Comment expliquer les disparités salariales », in *Les salaires en France*, 2006, INSEE : http://www.insee.fr/fr/ffc/docs_ffc/SALFRA06ac.PDF

Roxane Silberman et Irène Fournier, « Jeunes issus de l'immigration. Une pénalité à l'embauche qui perdure », *Bref*, n° 226, 2006, CEREQ : <http://www.cereq.fr/pdf/b226.pdf>

Patrick Simon et Martin Clément, « Rapport de l'enquête "Mesure de la diversité". Une enquête expérimentale pour caractériser l'origine », *Documents de travail*, INED, n° 139, 2006 : www.ined.fr/fr/ressources_documentation/publications/documents_travail/

Patrick Simon et Martin Clément, « Comment décrire la diversité des origines en France ? Une enquête exploratoire sur les perceptions des salariés et des étudiants », *Population & Sociétés*, n°425, 2006, INED : www.ined.fr/fr/ressources_documentation/publications/

TIES, Enquête internationale sur « L'Intégration de la Deuxième Génération en Europe » : <http://www.tiesproject.eu/index.php?lang=fr>,

Virginie Guiraudon, Karen Phalet et Jessika Ter Wal, « Rapport final Pays-Bas », in *Etude comparative de la collecte de données visant à mesurer l'étendue et l'impact de la discrimination dans certains pays*, Economie et Humanisme, 2004, Commission européenne : http://ec.europa.eu/employment_social/fundamental_rights/pdf/aneval/pb_fr.pdf