

Mesure des inégalités ou des discriminations : apports et limites des enquêtes

RIANDEY Benoît

Si la mesure des inégalités est un objectif ancien de la statistique sociale, celle des discriminations s'est imposée tardivement en France à la fin de la décennie 90 sous l'impulsion des directives européennes.

Cette distinction entre inégalités et discriminations est-elle technique ou au contraire juridique ou morale ? S'agit-il d'un nouvel avatar du débat entre égalité et équité ? La mesure des discriminations relève-t-elle du fait d'opinion ou d'une évaluation méthodologiquement pointue et incontestable ? Doit-on distinguer l'exercice de la preuve et la mesure de son amplitude ? Sur les sujets sensibles, les obligations juridiques de confidentialité ou les réserves psychologiques sont-elles des obstacles sans solution pour cette mesure ? Quels dispositifs pourraient y remédier ?

Les réponses à ces questions conditionnent l'évaluation de l'efficacité des politiques sociales, qu'elles soient tournées vers l'éducation, la santé, le logement ou l'emploi.

Débat autour de la notion de discrimination

L'intuition élémentaire de discrimination est suggérée par le constat qu'une sous-population ne bénéficie pas d'un dispositif, d'un avantage en égale proportion à sa représentation dans la population. Sa mise en évidence serait immédiate dès lors que le recensement dénombrerait cette sous-population tandis que les administrations et entreprises en dénombreraient pour leur part les bénéficiaires. Quelque instructif que soit ce ratio, il serait rapide d'interpréter les différences observées comme discriminatoires et d'ailleurs peut-on légitimement dénombrer toute sous-population quel qu'en soit le critère de définition : caractères physiques ou médicaux, origine géographique, appartenance religieuse, orientation sexuelle ? La notion de minorité invisible n'y fait-elle pas obstacle ?

De longue date, les démographes ont veillé à produire de saines mesures différentielles en s'écartant des proportions naïves et en construisant des taux échappant aux effets de structure. Il est impensable d'évoquer une surmortalité sans avoir tenu compte de la structure par âge des sous-populations concernées. Ces précautions techniques suffisent-elles pour s'élever des différences brutes à la mesure de discriminations ?

* INED.

Au sens des directives européennes, une discrimination directe interviendra quand une personne fera l'objet d'un traitement distinct pour un motif en cause (sexe, « race »...). Elle sera indirecte si une disposition juridique implique indirectement cette différence de traitement pour ce motif en cause.

Bien que l'obligation d'une mesure statistique relève d'une obligation juridique communautaire, statisticiens et juristes divergent dans leur définition, au profit d'un champ plus extensif de la part des statisticiens, contrôlé par les méthodes économétriques. La discrimination est alors définie comme l'inégalité résiduelle imputable à un critère une fois éliminés les effets de tous les facteurs légitimes d'inégalité. Ainsi la discrimination indirecte au sens du statisticien ne ressort pas nécessairement d'une disposition juridique. Elle est simplement statistiquement constatée sur un critère, indépendamment de toute volonté ou cause et bien souvent à l'insu des acteurs ; le statisticien admet donc l'existence de discriminations en l'absence de victimes individualisées et de discriminateur. Une caractéristique pourrait être à l'origine d'une discrimination à l'embauche sans qu'aucun recruteur n'ait perçu le désavantage octroyé aux détenteurs de ce caractère, ni qu'aucun de ses détenteurs n'y voit la cause de son échec à l'embauche. On se rapproche de la notion épidémiologique de facteur de risque, après contrôle de tous les facteurs légitimes de confusion. En matière d'embauche, tous ceux liés à la productivité ou l'adéquation au poste.

De plus, une inégalité peut résulter d'une discrimination antérieure : une discrimination en matière d'éducation induit des inégalités en matière de qualification et donc d'emploi. On pourrait parler de discriminations en chaîne. Le débat sur la parité est empreint de cette situation et politiquement s'intéresser à distinguer s'il s'agit d'inégalité ou de discrimination peut apparaître une question oiseuse face à cette exigence sociale.

La discrimination peut être imputable à un critère croisé tel que l'origine et le quartier sensible de résidence, et on parlera de discrimination multiple (Kachoukh, 2011). Mais les économistes s'efforceront de mesurer l'effet propre de chacun de ces facteurs de risque, par exemple l'effet du quartier de résidence ou au contraire celui de l'origine (Petit, 2010).

Le résultat prête à objections et controverses, ne serait-ce que pour l'effet supposé d'un facteur non pris en compte ou non mesuré. Néanmoins l'amplitude des discriminations balaie souvent ces critiques.

Sur la base de méthodes épidémiologiques inspirées des enquêtes cas – témoins, on met en œuvre deux plans d'expérience efficaces pour prouver, tester, l'existence des discriminations (Duguet, 2007) sans toujours pouvoir évaluer leur niveau. À l'AIDELF, on évitera de les appeler « testing ». L'objectif est toujours de mettre en évidence une discrimination directe, ségrégative, ou raciste.

D'un point de vue méthodologique, constituer deux lots d'individus appariés comme des faux jumeaux ne se distinguant que par un caractère (le sexe ou l'origine) est particulièrement efficace pour tester l'existence d'une différence de traitement. Mais l'absence de représentativité de chacun des deux échantillons par rapport à leur population n'augure aucunement que la mesure quantitative de l'écart, certes significatif, soit valide.

Si la méthode est très pertinente pour mettre en cause l'accès raciste à une boîte de nuit ou l'accueil discriminatoire des CV nominatifs, le recours à ces méthodes pour l'évaluation des politiques publiques reste à imaginer. On est davantage dans l'ordre du contentieux que de la statistique.

En définitive, la mesure de discriminations est un exercice statistique difficile. Si les statisticiens doutent, ne suffit-il pas de demander l'avis du public ou aux victimes (Beauchemin, 2010) ?

La perception des discriminations

Souvent les publications grand public confondent la mesure d'une discrimination avec la mesure de la perception d'une discrimination par le public ou la perception par les intéressés eux-mêmes d'une discrimination subie (Algava, 2006).

Les opinions mêmes relatives au classement par fréquence des motifs de discriminations seront très influencées par la visibilité comparée de ces délits et leur médiatisation. Il y a risque de surinterprétation de ces variables d'opinion.

Forts de cette critique, on peut interroger le public sur les actes de discrimination dont ils ont été témoins. C'est en particulier la revendication d'associations militantes qui dénie l'utilité d'enquêtes statistiques traitant des discriminations au motif qu'elles sont bien informées par les témoignages. Il s'agit donc d'une statistique de témoignages sur des actes délictuels et non directement d'actes délictuels.

Les démographes, experts en enquêtes indirectes sur l'émigration (Somoza, 1984) connaissent l'erreur d'estimation qui en résulte : pour passer de l'univers statistique des témoignages à ceux des faits, il faut pondérer chaque fait témoigné par l'inverse du nombre de témoins de ce fait. Sinon, on gît dans le biais d'un sondage à probabilités inégales dont on ignorerait les probabilités d'inclusion. Dans l'enquête de Barbade sur l'émigration, l'observation de chaque émigré cité par sa fratrie doit être pondérée par l'inverse de la taille de la fratrie. Autant dans une enquête, il est facile de connaître la taille de la fratrie d'un sujet, autant il est impossible de connaître le nombre de témoins d'un délit. Par ailleurs, il y aura toujours à corriger le biais des délits sans témoins, les crimes parfaits, comme celui des émigrés enfants uniques.

Reste alors la solution de demander aux intéressés de déclarer les discriminations dont ils estiment avoir été l'objet. C'est l'enquête de victimation. Certes la perception des discriminations subies est très corrélée avec les faits subis, même si l'auto-perception n'est pas la mesure factuelle. Les discriminations subies par un groupe peuvent ne pas être ressenties par les personnes concernées. La distinction faite entre bilan factuel et perception n'enlève aucune légitimité à la mesure de ces perceptions à condition de les considérer comme telles et de contrôler l'éventuel effet inducteur du questionnement.

Critères sensibles de discrimination

Autre difficulté méthodologique, les critères sensibles de discrimination peuvent être difficiles à collecter, tant pour des raisons juridiques que psychologiques, par les administrations, les recensements ou les enquêtes. La première occasion au cours duquel j'ai rencontré cette difficulté était un colloque au cours de laquelle la Caisse Nationale d'Allocations Familiales exprimait le dilemme suivant : « Nous pensons

que certaines familles étrangères ne perçoivent pas toutes les allocations auxquelles elles ont droit, dont l'allocation logement ; mais, compte tenu de l'universalité du droit aux allocations, la nationalité n'est pas un critère d'attribution ; sa collecte *excessive* pour la finalité de versement des allocations est illégale car contraire à la loi Informatique et Libertés ». Ainsi l'orateur estimait que seule une enquête par sondage pouvait permettre d'estimer la fréquence de ce phénomène (espéré bien rare), sans bien sûr permettre d'y remédier. Certes les directives européennes de 1990 relative aux discriminations et de 1995 relative à la protection de la vie privée ont depuis imposé ou autorisé ces traitements statistiques ; il n'en demeure pas moins qu'on ne peut que constater l'impuissance et le manque d'imagination de la statistique administrative en la matière. Les épidémiologistes, fidèles à Hypocrate, savent trouver une procédure statistique et un mode sécurisé de levée de l'anonymat si le sujet fait l'objet d'une pathologie précédemment inconnue exigeant urgemment des soins.

Plus fondamentalement que pour un obstacle juridique, c'est la sensibilité psychologique de l'information qui justifie de mettre en place des procédures sécurisées permettant de collecter et d'apparier des données standard administratives ou d'enquête avec des données personnelles sensibles. La loi permet le traitement statistique des données administratives mais n'interdit pas au statisticien de collecter, éventuellement exhaustivement, des données personnelles sensibles auxquelles l'administration ne doit avoir accès et de les apparier en instances sécurisées aux données administratives, donc en dehors du site de l'administration. Encore faut-il le vouloir et savoir le faire. Les techniques d'appariements sécurisés sont disponibles à cet effet (Quantin, 2005). La question à résoudre est d'abord celle de la confiance.

Nous en avons proposé une solution pour permettre aux grandes entreprises de tester l'absence de discrimination de carrière (après l'embauche) par rapport à des critères auxquels les services des ressources humaines n'ont pas le droit d'accéder (Quantin, 2009). Le principe en est de demander au sujet de répondre à un questionnaire auto-administré sur ordinateur portable ou sur internet, questionnaire aussitôt crypté, puis apparié à l'extérieur à des données administratives anonymisées le concernant grâce à un identifiant définitivement anonymisé par hachage SHA (cf. annexe). Du point de vue de la sécurité et de la légalité, la procédure est solide. Reste à mobiliser la confiance par une intermédiation reconnue : syndicats, délégués du personnel, CNIL¹, ex-Halde².

Cette méthode pourrait être utilisée par Pôle emploi pour mesurer globalement les discriminations à l'embauche. Au cours d'une simple enquête statistique, l'appariement est immédiat : il suffit de passer l'ordinateur portable muni d'un casque à l'enquêté après avoir montré le résultat d'un cryptage d'informations. L'enquête Violence et santé de la DREES en fut le premier exemple (Cavalin, 2008) avant d'être repris par l'enquête de victimation de l'ONDRP, l'observatoire national de la délinquance.

Notre proposition a fait l'objet d'une révision simplifiée dont les limites sont conceptuellement intéressantes : Nous avons proposé qu'à l'embauche, le nouveau salarié réponde confidentiellement à un questionnaire sensible aussitôt crypté, qu'il saisisse son numéro Nir de sécurité sociale avec sa clé de contrôle aussitôt haché et que ce fichier anonymisé et crypté soit stocké par un hébergeur de

¹ Commission nationale de l'informatique et des libertés.

² Mais la disparition de la *Haute autorité de lutte contre les discriminations et pour l'égalité*, fondue par le parlement dans les missions d'un défenseur des libertés semble avoir bien effacé la volonté collective de lutter contre les discriminations en même temps que cette instance de référence.

données sensibles. À la demande de l'employeur, cet hébergeur recevrait un extrait du fichier historique du personnel dont les enregistrements individuels seraient identifiés par leur numéro Nir haché avec la même clé en sorte de récupérer le même identifiant anonyme que celui détenu dès l'embauche par l'hébergeur. Ainsi, longtemps après l'embauche, pourrait-on apparier les caractéristiques sensibles initiales du salarié avec son déroulement de carrière. Bien sûr un questionnaire de mise à jour pourrait être réalisé à espacement régulier.

Une grande simplification a été proposée, celle de ne faire remplir le questionnaire sensible par le salarié qu'à la veille du traitement statistique, juste après lui avoir remis son extrait de carrière. Le salarié se contente donc de compléter l'enregistrement fourni, ce qui dispense de l'appariement, de l'usage juridiquement complexe du Nir et du hachage d'un identifiant devenu parfaitement superflu. Pourquoi faire compliqué quand on peut faire simple ?

Le prix à payer de cette simplification est celui de l'abandon d'un échantillon longitudinal au profit d'un échantillon rétrospectif et donc le prix du biais d'échantillonnage bien connu des enquêtes rétrospectives qui en résulte : les salariés qui ont quitté l'entreprise à la date du traitement statistique sont exclus de l'analyse. Celle-ci répond à la question : « *Des salariés actuels ont-ils fait l'objet de discrimination de carrière depuis leur embauche ?* » Même réalisée régulièrement, cette question ne peut être assimilée à la question posée « *Des salariés ont-ils fait l'objet d'une discrimination de carrière depuis leur embauche ?* » sauf à supposer que les sorties de l'entreprise soient indépendantes des faits discriminatoires qui s'y produisent. Ce serait supposer qu'un salarié victime de discrimination n'aurait pas une plus forte propension au licenciement, à la démission ou au suicide. Une hypothèse que le traitement statistique viserait fondamentalement à tester... et donc pas à prendre comme une certitude préalable (Riandey, 2010).

Cette hypothèse a d'ailleurs été infirmée par les travaux de Thierry Laurent et Ferhat Mihoubi (Laurent, 2010) relatifs à l'effet de l'orientation sexuelle sur les discriminations salariales. Rappelons que les auteurs identifiant dans l'enquête Emploi de l'INSEE des paires possiblement en couple homosexuel et comparaient leurs salaires, toutes choses égales par ailleurs, à ceux des autres salariés. À l'égard des hommes, ils observaient une absence de différence pour la période proche de l'embauche, puis une discrimination négative qui disparaissait ultérieurement. Cette disparition à la longue résultait selon eux d'une sortie de l'entreprise, tandis que la première période intervenait avant la révélation de l'homosexualité. Les femmes homosexuelles bénéficiaient au contraire d'une discrimination positive, sans doute l'exonération d'une discrimination des salariées plus fortement soumises au « risque de maternité ». De la sorte d'ailleurs, ils réalisaient la mesure la plus convaincante de la mesure de la discrimination salariale que subissent les salariées compte tenu des rôles familiaux en France.

La mesure statistique de la discrimination à l'embauche par une entreprise est beaucoup plus délicate à opérer et elle a la réputation d'être la plus massive, en particulier à l'égard des étrangers ou des personnes d'origine étrangère.

Les autres travaux des économistes d'Evry (Duguet, 2007) sont intéressants à cet égard. Leurs réponses par paire dissimulée de cv aux annonces d'embauche constituent des tests très convaincants des discriminations liées à un nom ou prénom étranger... L'exemple de la profession d'analystes informaticiens est probant à cet égard. Ils ne peuvent toutefois pas constituer au sens statistique une estimation sans

biais de la probabilité différentielle de réponse à un cv due à ces caractères. Mais mettre en évidence un phénomène massif fait passer au second plan l'évaluation du biais.

Ont surpris, au contraire, les résultats négatifs de Thomas Le Barbanchon (2012) au CREST montrant l'inefficacité des cv anonymes pour neutraliser les discriminations à la prise de rendez-vous pour un entretien d'embauche. Le débat n'est sans doute pas clos.

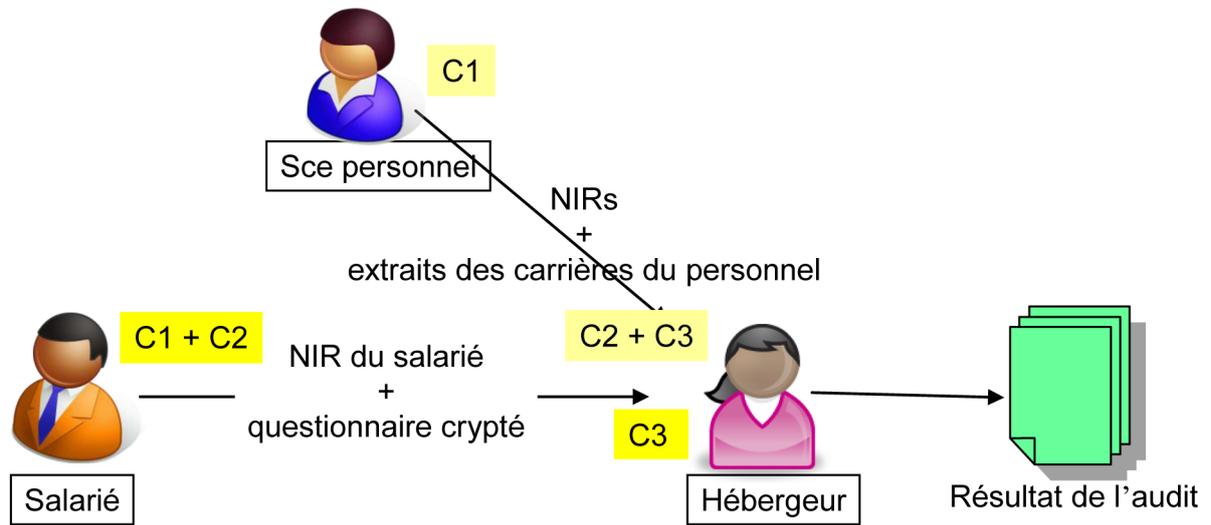
En définitive, les test(ing)s demeurent les mesures les plus incontestées des discriminations alors qu'elles ont une forte valeur de preuve, mais la faiblesse statistique de mesures des dispositifs juridiques. La mesure économétrique des discriminations est la voie scientifique la plus valide, mais toujours susceptible de mise en cause à l'évocation de la variable omise ou non observée. Indemnes de cette critique, la mesure des inégalités et celle de la perception des discriminations subies conservent donc leur importance majeure.

Références bibliographiques

- Algava E, Beque M., 2006, « Perception et vécu des comportements intolérants. Une analyse du module *Relations avec les autres* de l'enquête Histoire de vie ». *Économie et Statistique*, n° 394.
- Beauchemin C., Hamel C., Lesne M., Simon P., 2010, « Les discriminations : une question de minorités visibles ». *Population et Sociétés*, n° 466.
- Cavalin C., 2008, « *Événements de vie et santé : principe et méthodes* » dans Guilbert et all. Eds, *Méthodes de sondage*. Actes du colloque Sondages2007 de la SFdS, Dunod : 108-1112.
- Duguet E., Leandri N., L'Horty Y., Petit P., 2007, « *Discriminations à l'embauche : un testing sur les jeunes de banlieues d'Île de France* ». Rapport pour le Centre d'Analyse Stratégique.
- Kachoukh F., Maguer A., Marnas A., 2011, « La discrimination multicritère à l'encontre des femmes immigrées ou issues de l'immigration sur le marché du travail », Halde.
- Laurent T., 2010, « Moins égaux que les autres ? Discrimination salariale en France ». Séminaire du Lundi. Séance du 10 mai 2010. INED.
- Le Barbanchon T., 2012, « Incidence du cv anonyme ». Séminaire du Lundi. Séance du 30 janvier 2012. INED
- Petit P., 2010, « Inégalités d'accès aux emplois et inégalités de salaires : la situation des enfants d'immigrés et des jeunes résidant dans une zone urbaine sensible ». *Économie et Statistique*, n° 433-434.
- Quantin C., Gouyon B., Allaert F.A., Cohen O., 2005, « Méthodologie pour le chaînage de données sensibles tout en respectant l'anonymat : application au suivi des informations médicales ». *Le Courrier des statistiques*, n° 113-114.
- Quantin C., Riandey B, Cohen O., Benoit A.M., 2009, « *Proposition à l'intention de la COMEDD d'un dispositif sécurisé de suivi statistique de la diversité dans les entreprises à l'attention des entreprises signataires de la Charte de la diversité* ». in HERAN F. Rapport de la Commed.
- Riandey B., Benoit A.M., Cohen O., Quantin C., 2010, « *Dispositif sécurisé de suivi statistique sensible dans les entreprises ou les administrations* », XVII^{es} Journées du Longitudinal, CEREQ.
- Somoza J. et al., 1984, « Barbados experimental migration survey : methods and results », IUSSP newsletter n° 20, 33-124.

Annexe

Audit des déroulements de carrières



Date d'embauche Date de l'audit

Clés de hachage du NIR
Clés de hachage du NIR