

Annexe 1 : L'origine des immigrés et de leurs enfants dans les enquêtes de la statistique publique. Quels principes déontologiques ?

Synthèse de l'exposé de François Héran

à la réunion du Conseil scientifique d'Économie et Statistique du 28/9/2005

En matière de questionnement et d'exploitations statistiques sur les origines nationales et ethniques, l'Insee applique depuis une dizaine d'années une doctrine, dont les grands principes avaient été synthétisés pour la première fois dans l'article de F. Héran du rapport du Commissariat Général du Plan *Immigration, marché du travail, intégration*, La documentation Française, 2002, p. 121-133¹. Il s'agit de rappeler les principes déontologiques qui peuvent guider le statisticien compte tenu, d'une part, de l'évolution des pratiques observées depuis une quinzaine d'années au sein de la statistique publique et, d'autre part, des avis rendus par la Commission nationale de l'informatique et des libertés.

Avant d'énoncer cette doctrine, il importe d'opérer une première distinction en amont : cette doctrine ne concerne pas les fichiers administratifs ayant pour objet de gérer des dossiers individuels mais uniquement les fichiers statistiques établis à des fins de connaissance, sans incidence sur le sort des personnes, sachant que ces fichiers sont généralement tirés d'échantillons aléatoires et traités sur un mode anonyme (même si l'enquête commence toujours par une phase de collecte nominative, pour des raisons de contrôle du terrain ou de suivi longitudinal).

Une autre distinction doit être pratiquée en aval. Les possibilités ou les restrictions présentées ici portent sur la collecte des données et leur traitement. D'autres restrictions, par ailleurs, s'appliquent aux phases de diffusion et portent sur le degré de détail des nomenclatures de publication. Il pourrait se faire, en effet, que certains tableaux de données publiés à un niveau de détail très fin permettent indirectement de lever l'anonymat de telle ou telle personne et de la retrouver (par exemple, en isolant le représentant unique d'une profession dans une petite commune et en croisant l'information ainsi obtenue avec un annuaire professionnel). Il existe des normes de diffusion à ce sujet, édictées par la Cnil pour les tableaux du recensement ou par l'Insee pour les résultats d'autres enquêtes. Dans la pratique, *Économie et Statistique* n'est guère concerné par ces restrictions, car les tableaux ou les analyses de régression publiées dans la revue opèrent à un niveau d'agrégation qui échappe d'entrée de jeu à ces restrictions. Du moins en est-il ainsi pour les ménages. L'anonymat des entreprises dans des enquêtes de branche pose évidemment de tout autres problèmes, qu'on n'abordera pas ici.

En quoi consiste la doctrine relative aux variables décrivant les origines des immigrés ou des personnes issues de l'immigration ? Elle revient à distinguer, comme le montre le tableau ci-après, trois niveaux d'information, tout en indiquant pour chacun (i) les possibilités offertes aux statisticiens, possibilités bien plus larges qu'on ne croit, et (ii) les précautions qui s'imposent pour se conformer à la jurisprudence sur le traitement des données individuelles.

¹ François Héran, « Les immigrés et leurs descendants dans le système statistique français : quelques réflexions sur les pratiques et les principes », in *Immigration, marché du travail, intégration, Rapport du séminaire présidé par F. Héran*, Commissariat Général du Plan, Paris, La documentation Française, 2002, p. 121-133.

Rappelons que, contrairement à une idée reçue, il est permis depuis longtemps de poser des questions « sensibles » dans une enquête de la statistique publique, c'est-à-dire une question risquant de faire apparaître directement ou indirectement des appartenances (réelles ou supposées) à un parti politique, un syndicat, une religion, une race, une orientation sexuelle. Mais poser de telles questions n'était possible avant la loi d'août 2004 qu'à une double condition : que la question soit pertinente pour le questionnaire (proportionnée aux objectifs poursuivis) et que les répondants donnent leur accord exprès, c'est-à-dire signé (la signature étant apposée non pas sur le questionnaire, qui perdrait ainsi son caractère anonyme, mais sur une feuille à part présentée par l'enquêteur). Cette procédure exceptionnelle de recueil du consentement exprès avait été adoptée en 1992 pour l'enquête sur la mobilité et l'insertion sociale des immigrés (enquête *MGIS*) non pas en raison des questions sur les origines nationales mais parce que certaines questions (sur l'observance du Ramadan et le lieu d'enterrement) faisaient apparaître l'affiliation religieuse.

Des expériences évoquées dans ce tableau, il ressort que la Commission nationale de l'informatique et des libertés n'a jamais considéré comme sensibles – et, par conséquent, n'a jamais assorti d'une obligation de recueil du consentement exprès – les questions d'état civil relatives au pays de naissance des répondants ou aux pays de naissance de leurs parents. En témoigne notamment l'enquête *Famille* de 1999, qui constitue l'enquête la plus officielle qu'on puisse imaginer (près de 400 000 personnes sollicitées par l'Insee lors du recensement) : on y trouve des questions sur le pays de naissance de la personne interrogée, les pays de naissance de ses deux parents, les langues parlées dans l'enfance, – autant d'informations qui, pour la Cnil, relevaient des appartenances culturelles ou nationales et non pas raciales ou ethniques. Pour autant, de telles questions ne sont pas anodines puisqu'elles retracent l'histoire familiale. Aussi a-t-il été décidé, d'un commun accord entre l'Insee, le Cnis et la Cnil, que l'enquête *Famille* demeurait facultative, alors que le recensement restait obligatoire.

Il importe également de rappeler – c'est un point-clef de la doctrine – que, contrairement à une idée répandue, la statistique publique travaillant à des fins de connaissance et non de gestion est parfaitement habilitée à « faire des différences entre les Français en fonction de leurs origines » et cela de très longue date : l'indicateur de nationalité, qui permet de savoir si l'on est Français de naissance ou Français par acquisition, figure dans tous les recensements depuis 1871. De même, le recensement permet de connaître le détail de la nationalité antérieure des immigrés devenus français : elle est demandée à chaque recensement depuis 1962. Il n'est donc pas contraire à la tradition républicaine de différencier les Français selon l'origine (ou, réciproquement, de différencier les immigrés selon qu'ils ont acquis ou non la nationalité française).

Un dernier élément de la doctrine appliquée par l'Insee en matière de statistiques des origines a trait à la déontologie de l'exploitation statistique. Sous peine d'imputer faussement les différences observées à des oppositions d'essence entre groupes ethniques, les données sur les origines des immigrés ne peuvent servir isolément de variables descriptives ou explicatives. Les comparaisons entre groupes doivent se faire à durée de séjour égale (les groupes étant d'abord séparés par des écarts de temps) et, dans toute la mesure du possible, après avoir contrôlé les autres variables qui contribuent à la construction des différences, qu'elles soient démographiques, sociales ou économiques. On fera de même pour diagnostiquer les situations de discrimination, qui ne se résument pas au constat des différences ou des inégalités

brutes mais qui doivent persister « toutes choses égales d'ailleurs » et revêtir un caractère systématique. Certes, l'exigence d'une analyse multivariée s'impose dans tous les secteurs de la statistique sociale ; elle constitue aujourd'hui un point d'honneur minimal pour un statisticien digne de ce nom, mais cette exigence prend une valeur déontologique particulière en présence de variables aussi chargées que les variables d'origine. C'est particulièrement à ce stade, dans l'administration de la preuve et dans sa présentation, que la rédaction d'*Économie et Statistique* se doit d'être vigilante. Cette vigilance doit aussi porter en amont sur la spécification du modèle (l'omission de variables-clefs pour mener une analyse « toutes choses égales d'ailleurs » pouvant sérieusement biaiser les résultats)

La doctrine de l'Insee en matière d'origines est-elle susceptible d'évoluer ? Non seulement elle entérine l'évolution des quinze dernières années (comme la jurisprudence de la Cnil pour des enquêtes majeures), mais elle permet aujourd'hui même de faire évoluer les pratiques actuelles en connaissance de cause. C'est en menant une réflexion approfondie sur la finalité des enquêtes et en cherchant à définir des variables plus appropriées aux nouveaux objectifs que l'Insee a pu récemment faire glisser plusieurs enquêtes du premier degré de connaissance vers le deuxième degré. Des questionnaires classés auparavant comme généralistes et qui, de ce fait, se contentaient de distinguer les immigrés des natifs, ont été enrichis de façon à permettre le repérage des enfants d'immigrés. C'est le cas de l'enquête *Formation et Qualification Professionnelle* (FQP) en 2003 et de l'enquête *Emploi* depuis janvier 2005, qui rejoignent ainsi l'enquête *Famille* de 1999, sans compter les opérations menées par d'autres organismes comme les enquêtes *Génération*s du Cereq... Il devient désormais possible de mesurer les inégalités et, le cas échéant, les discriminations qui séparent les descendants d'immigrés des descendants de natifs, en particulier sur le marché du travail.

Dans la hiérarchie des degrés de connaissance sur le rôle des origines, l'enquête *Trajectoires et origines* » actuellement en préparation pour faire suite à l'enquête *MGIS* de 1992, se situe sans ambiguïté au troisième et dernier degré. Elle a l'ambition de récolter non seulement des informations sur les pays d'origine et les langues parlées, mais aussi sur les appartenances ethniques déclarées ainsi que sur les principales qualités perceptibles qui peuvent servir de support aux discriminations dans notre société : couleur de la peau, coiffure, tenue vestimentaire, accent et autres signes corrélés de façon visible ou hypothétique à une appartenance religieuse ou ethnique (pratiques alimentaires, respect d'un calendrier festif non chrétien, pratiques funéraires, etc.). Elle devra permettre également aux intéressés de relater les incidents de parcours et d'apprécier eux-mêmes le rôle qu'ont pu y jouer ces qualités ou ces signes. Le recueil de ces données est licite, à condition qu'elles soient pertinentes pour l'enquête, qu'elles obtiennent l'assentiment éclairé des enquêtés², ce qui peut se faire en recueillant leur accord écrit et/ou en rendant l'enquête facultative. Il fait partie de la mission de la rédaction d'*Économie et Statistique* de vérifier que ces protocoles ont été respectés et d'en faire état dans la publication, par exemple sous la forme d'un encadré.

² Pour l'Insee ou les SSM, ce consentement n'est plus nécessaire à partir du moment où la Cnil a donné son autorisation.

Tableau 1

Graduation des types d'information qu'il est possible de recueillir actuellement dans les enquêtes anonymes de la statistique publique, avec les obligations correspondantes

<i>a</i>	Type d'information (chaque degré ajoutant de nouvelles informations au degré précédent)	1er degré Pays de naissance Indicateur de nationalité (<i>Fr. de naissance / devenu Fr. / étranger</i>) Année d'entrée en France Nationalité actuelle Nationalité antérieure	2e degré Idem + ... Pays de naissance de la mère et du père Nationalité antérieure de la mère et du père Langues reçues des parents Attachement à la religion (<i>sans mention de l'affiliation</i>)	3^e degré Idem + ... Origines des grands-parents Distinction entre enfant de rapatriés et enfant d'immigrés algériens Apparence physique (dont couleur de la peau) Appartenance ethnique (<i>infra- ou transnationale : Kurde, Kabyle, Rom...</i>) Affiliation religieuse Etc.
<i>b</i>	Génération concernée	Ego (y c. conjoint)	Ego + parents et beaux-parents d'Ego	Toutes générations
<i>c</i>	Types d'enquête où l'information est pertinente	Recensement + enquêtes généralistes standard	Enquêtes spécialisées sur la mobilité sociale, professionnelle, résidentielle, la transmission de l'éducation ou des langues, la construction des identités	Enquêtes spécialisées sur la discrimination à raison de la couleur, de la religion, de l'orientation sexuelle, etc.
<i>d</i>	Exemples d'enquêtes ayant posé ces questions	Enq. <i>Emploi</i> Enq. <i>Logement</i> Autres enquêtes générales	Enq. <i>Famille</i> associée au RP199 (EHF) [n = 380 000] Enquête <i>FQP</i> Enq. <i>Éducation</i> Enq. <i>Histoire de vie</i> (ou <i>Construction des identités</i>) Enq. <i>Emploi</i> rénovée en 2005	Enq. <i>MGIS</i> (uniquement en raison des questions sur la religion)
<i>e</i>	Question "sensible" au sens des lois de 1978 et 2004	Non	Non	Oui
<i>f</i>	Consentement écrit nécessaire ?	Non	Non	Oui
<i>g</i>	Enquête obligatoire (décidé par le Comité du Label)	Obligation possible	Facultative	Facultative