

Quantifier les discriminations raciales entre données d'enquête et données expérimentales : l'évolution du rôle du chercheur dans la construction de la variable « race »

Cléo Chassonnery-Zaïgouche
Centre Walras-Pareto, Université de Lausanne

Léontine Goldzahl
LEDA-LEGOS, Université Paris-Dauphine

Charlotte Levionnois
Centre d'économie de la Sorbonne, Université Paris 1 Panthéon-Sorbonne et
OCDE

Dans la période récente, la micro-économie appliquée a été fortement renouvelée par l'introduction de méthodes expérimentales, qu'il s'agisse d'expériences de terrain ou de laboratoire. Ces méthodes ont permis de produire de nouveaux types de données qui sont de plus en plus utilisées en économie, aux côtés ou en substitution de données quantitatives issues d'enquêtes. L'introduction de ces nouvelles méthodes a modifié le processus de production des données, et a, en particulier, accru le rôle du chercheur dans leur production.

Ce chapitre se concentre sur un cas emblématique de ce changement : les données produites dans le but de mesurer les discriminations raciales. L'analyse des discriminations constitue un exemple particulièrement adapté car, depuis les années 1970, différentes méthodes microéconomiques ont été utilisées pour l'étudier, et la plupart des articles sur la question ont été publiés dans des revues d'économie représentatives des évolutions méthodologiques de la discipline.

Dans ce chapitre, nous appellerons « données » des informations sur des individus rassemblés dans un échantillon, représentatif ou non d'une population, informations qui sont dénombrées, mises en série et utilisées pour quantifier des phénomènes (sociaux, économiques, démographiques). Nous appellerons « donnée primaire » l'information elle-même. Les données primaires concernent des caractéristiques extrinsèques (informations socio-économiques comme le revenu par exemple) et intrinsèques (le sexe ou

l'origine ethnique), qu'elles aient ou non une existence juridique. Ce dernier élément est central : il détermine en partie la marge de manœuvre du chercheur. À partir des données, qu'elles soient produites ou non par le chercheur lui-même, celui-ci construit des variables¹.

Deux types majeurs de données quantitatives ont été utilisées pour mesurer les discriminations raciales : les données issues d'enquête, parmi lesquelles figurent les données issues du recensement, d'une part, et les données expérimentales d'autre part. Les données d'enquête sont obtenues à partir des réponses à un questionnaire soumis à un échantillon d'individus issus d'une population cible. Les données de recensement sont des données d'enquête spécifiques car le recensement vise l'exhaustivité, la population cible est en général la population entière d'une entité géographique juridique, comme un État. Les données expérimentales correspondent aux données issues de la collecte des résultats de décisions prises par des individus dans le cadre d'une expérience en laboratoire ou sur le terrain. Ces décisions sont encadrées par un protocole défini par le chercheur.

Faire « l'inventaire des acteurs » impliqués dans la production de ces deux types de données permet d'étudier la transformation du rôle du chercheur dans ce processus. Soulignons que les discriminations raciales elles-mêmes ne sont pas l'objet de l'analyse, mais que ce sont les méthodes de production des données, et en particulier l'usage de la catégorie raciale qui nous intéressent ici. L'analyse des discriminations liées à l'origine des individus pose des problèmes méthodologiques particuliers. Ce qui apparaît comme deux objets distincts – la production de données dans le but de mesurer les discriminations et les catégories raciales – ne vont pas l'un sans l'autre, historiquement et conceptuellement. Mesurer la discrimination consiste à mesurer l'effet de l'usage de cette catégorie par les acteurs. Sans l'existence de catégories raciales, le chercheur ne peut mesurer le phénomène discriminatoire, qu'il s'agisse de l'impact de la discrimination

¹ Une variable statistique (on parle aussi de caractère statistique), est ce qui est observé ou mesuré sur chaque individu d'une population statistique. Une variable statistique peut prendre différentes modalités ou valeurs. Ces modalités peuvent être une grandeur susceptible d'être mesurée par un chiffre (revenu en euros, nombre d'enfants) ou une qualité susceptible d'avoir plusieurs modalités possibles (ex. H/F, ou statut matrimonial : célibataire, concubin, Pacsé, marié, etc.) Une variable statistique définit donc une partition d'une population déterminée ; chaque individu étant caractérisé par une et une seule modalité pour ce qui concerne chaque qualité, et chaque individu étant caractérisé par un seul chiffre pour les variables quantifiées. On peut définir potentiellement une multitude de variables statistiques pour chaque individu de la population observée. Il y a autant de partitions de la population que de variables statistiques retenues.

sur des grandeurs économiques (e.g. discrimination salariale) ou de la mise en évidence de comportement discriminatoire au sein d'un laboratoire.

D'où l'importance d'un « détour historique » retraçant l'évolution de ces catégories : en effet, pour le chercheur en économie, la donnée raciale a d'abord été construite par d'autres acteurs – notamment administratifs, et considérée comme exogène dans le cadre de son analyse. Avec l'émergence des méthodes expérimentales en revanche, le chercheur a désormais la possibilité de construire la variable « race² » au-delà de la catégorie légale du recensement.

La première section est consacrée à l'histoire des multiples acteurs de la production de données d'enquête largement utilisées pour mesurer la discrimination raciale aux États-Unis. Dans une seconde section, nous revenons sur la production par le chercheur des données et des catégories raciales dans le cas de données expérimentales. L'objectif de ce chapitre est de montrer le changement du rôle du chercheur dans la production des données et sa principale conséquence, la modification des possibilités de recherche.

Utiliser des données d'enquête pour mesurer les discriminations

Deux exemples de bases de données largement utilisées pour l'étude de la discrimination : le CPS et le PSID

La *Current Population Survey* (CPS) est une base de données issue directement du recensement dont l'origine est administrative. Le *Panel Study of Income Dynamics* (PSID), né dans le sillage de la « Guerre contre la pauvreté » des années 1960, est géré par le *Survey Research Center* de l'Université du Michigan. Ces deux bases de données sont obtenues à partir d'enquêtes. Elles utilisent toutes les deux les catégories raciales du recensement et rassemblent d'autres informations socio-économiques. Ces deux bases de données ont été utilisées par les deux articles de références du

² Nous utilisons les termes « race » entre guillemets et le terme racial(e) sans guillemets. Du point de vue des sciences sociales, le terme correct est « racialisé » pour mettre en évidence l'aspect construit du processus de racialisation des individus (i.e. leur attribuer une « race »). Nous avons préféré reprendre le terme usité aux États-Unis pour des raisons pratiques. Dans ce pays, « la réfutation des fondements biologiques et anthropométriques des distinctions raciales [notamment par Franz Boas] n'a pas eu pour conséquence la contestation de l'usage de la race comme catégorie sociale » (Schor 2009: 16). Cet usage repose sur des bases conventionnelles.

corpus sur la mesure des discriminations en économie (Blinder 1973 ; Oaxaca 1973).

La Current Population Survey

La CPS représente un bon exemple de base de données issue directement du recensement, utilisée par exemple pour mesurer les inégalités dans l'accès à l'emploi. Enquête mensuelle conduite par des agents recenseurs (par opposition à auto-administrée), effectuée par le *Bureau of Labor Statistics*³ (BLS), elle a été créée initialement pour mesurer le taux de chômage⁴. Elle rassemble des informations concernant l'emploi – la « participation » au marché du travail, le temps de travail, les revenus et caractéristiques démographiques des individus – et permet d'identifier les individus en fonction de critères raciaux. L'enquête, composée de 60 000 ménages est continue et constituée de données en coupe instantanée, collectées tous les mois pendant quatre mois, interrompue ensuite pendant huit mois, puis de nouveau pendant quatre mois⁵. La structure de l'enquête est par ailleurs révisée lors des recensements décennaux. L'échantillon est représentatif de la population américaine civile, et à partir des années 1970, des sous-échantillons représentatifs par État sont fabriqués.

Ces données sont accessibles via le Census Bureau ou des plateformes d'accès liées à des centres de recherche – comme les Integrated Public Use Microdata Series (IPUMS) de l'Université du Minnesota ou le National Bureau for Economic Research (NBER). De manière générale, l'analyse de données et la production d'indicateurs, l'extension des bases, la refonte des enquêtes, le test des questionnaires, la collecte de données supplémentaires et la mise en forme des données est le travail du BLS. La CPS est un exemple de base de données surdéterminée de manière administrative. La production des données est caractérisée par les étapes d'une sorte de

³ La CPS est le prolongement d'une enquête par carte postale effectuée en 1937, pour laquelle les réponses aux questionnaires étaient inscrites sur une carte à renvoyer sans affranchissement. Son institutionnalisation date de mars 1940, au moment du recensement décennal. La CPS est gérée par une agence indépendante jusqu'en 1942, date à laquelle sa gestion est transférée au *Census Bureau*. En 1959, cette responsabilité est prise en charge par le BLS, rattaché à l'administration fédérale, en coopération avec le *Census Bureau* pour les suppléments statistiques. Les principales évolutions sont l'implémentation du concept d'activité qui définit les situations d'emploi, de chômage et d'inactivité, l'expansion de l'échantillon (de 8 000 ménages en 1940 à 23 000 en 1942, pour 60 000 ménages aujourd'hui) et la consolidation des données accessibles dans des recueils de données.

⁴ Pour une histoire officielle, voir Bregger (1984) et US Census et BLS (2006).

⁵ Il existe des « suppléments » constitués en dehors de ces temporalités, qui concernent des éléments en particulier ; par exemple, un supplément centré sur le marché du travail est ainsi collecté en mars de chaque année.

« modèle linéaire⁶ » de la production de la donnée. Les données sont produites par les administrations et mises en série, le chercheur est chargé de les analyser en utilisant les instruments de la science appliquée, ses résultats servant à éclairer des questions sociales et politiques. Dans ce processus, le chercheur utilise des données, notamment raciales, produites par d'autres acteurs et n'intervient que pour créer des variables synthétiques à partir des bases de données administratives et des variables fournies par l'administration.

Le Panel Study of Income Dynamics

Le PSID est aussi fortement associé à la mesure de la discrimination raciale aux États-Unis. Cette enquête présentait en effet l'avantage d'indiquer le salaire moyen pour l'année précédant l'enquête : c'est la présence de cette variable, ajoutée à d'autres d'ordre social, démographique et économique, qui permet d'identifier les écarts de salaire observables entre des individus et surtout, de quantifier la part de cet écart de salaire uniquement imputable à la variable « race », et non aux autres caractéristiques individuelles.

Le PSID, bien que de même nature que le CPS, i.e. comportant des données d'enquête, si il trouve son origine dans des programmes fédéraux, n'est plus géré de manière administrative. À la suite de la signature du Civil Act de 1964, l'administration Johnson mets en place l'Office of Economic Opportunity, chargée, en tant qu'agence fédérale, d'appliquer les principales mesures antidiscriminatoires⁷. Cette agence a supervisé la production de l'enquête sur les « opportunités économiques » (Survey of Economic Opportunity, SEO, 1967) qui deviendra, un an plus tard, l'enquête de panel des dynamiques de revenu (Panel Study of Income Dynamics, PSID, 1968) lors de son transfert à l'Université du Michigan (au Survey Research Center).

Le Panel Study of Income Dynamics (PSID) est une enquête longitudinale⁸ dont le champ comprend l'ensemble des familles américaines.

⁶ En histoire et sociologie des sciences, le modèle linéaire est un moment historique caractérisé par une relation linéaire entre sphère politique et sphère scientifique (Armatte 1995). Ce modèle « suppose que les sciences pures peuvent se prolonger par des sciences appliquées qui elles-mêmes peuvent s'illustrer dans des dispositifs techniques rencontrant – quel hasard! – des besoins sociaux » (Armatte 2010 : 11).

⁷ Sur l'histoire de la « guerre contre la pauvreté », voir O'Connor (2001).

⁸ Une enquête, ou plus généralement une mesure longitudinale, résulte du suivi d'une population dans le temps en fonction d'un évènement ou d'une caractéristique de départ. Ainsi, un individu est interrogé plusieurs fois sur les mêmes caractéristiques, ce qui permet essentiellement d'étudier le changement. Les termes de mesures répétées ou de données diachroniques sont parfois utilisés en français.

Conduite depuis 1968 dans les mêmes familles et leurs descendants, elle constitue la plus vieille enquête en panel au monde. Après avoir été interviewées pour la première fois en 1968, les familles l'ont été tous les ans jusqu'en 1997. À partir de 1997, les données ont été collectées tous les deux ans. Les individus qui quittent le foyer familial au cours du temps sont suivis lorsqu'ils forment un nouveau ménage et sont interrogés dans leur nouveau foyer. Pour éviter la non-représentativité du nouvel échantillon lié à la construction du panel, de nouveaux individus ont été interrogés, rééquilibrant ainsi la représentativité de la population américaine dans l'échantillon.

Les informations collectées lors des entretiens d'environ 75 minutes concernent une grande variété de sujets économiques, démographiques, sociaux, de santé, géographiques ou encore psychologiques⁹ sur l'ensemble du cycle de vie des différentes générations des familles considérées. Cette enquête rassemble les données qui décrivent les caractéristiques familiales dans leur ensemble, de même que les caractéristiques propres aux membres constituants de la famille considérée. Bien que certaines informations soient collectées pour tous les individus de la famille, le niveau de détail des informations est accru pour les adultes dits « à la tête du ménage » (appelés « head and wife »¹⁰).

La méthode d'entretien des individus a changé au cours du temps. Jusqu'en 1972, les entretiens sont réalisés lors d'un rendez-vous individuel en face à face, en utilisant un formulaire papier (cf. illustration 6) ; à partir de 1973, les entretiens sont effectués par téléphone. Depuis 1993, la procédure est assistée par ordinateur via le Computer Assisted Telephone (CAT). Parmi les sujets abordés aujourd'hui, de nombreux datent des premières enquêtes de 1968 mais une centaine d'autres variables thématiques ont été ajoutées lors des différentes restructurations de l'enquête. Enfin, les données sont anonymisées avant d'être transmises aux utilisateurs.

Une des forces du PSID repose sur son taux de réponse élevé, conséquence du faible taux d'attrition et du bon suivi des jeunes adultes

⁹ Plus précisément l'emploi et les salaires afférents, les revenus du capital et de transfert, les dépenses, le logement, l'éducation, la situation géographique, l'état de santé, les comportements de santé, l'assurance santé, l'historique de l'état de santé, le bien-être émotionnel, la satisfaction dans la vie, la mortalité et ses causes, le mariage et la fertilité, la participation aux programmes gouvernementaux, les difficultés financières, la propriété de véhicules, la richesse et les pensions, ainsi que des questions liées à la philanthropie.

¹⁰ "Within each wave of data, each family unit has one and only one current Head. Originally, if the family contained a husband-wife pair, the husband was arbitrarily designated the Head to conform with Census Bureau definitions in effect at the time the study began." (PSID FAQ, <https://psidonline.isr.umich.edu/Guide/FAQ.aspx?Type=5#270>). La catégorie a évidemment évolué au cours du temps, mais reste très fortement marquée par cette hypothèse patriarcale.

lorsqu'ils forment leur propre famille. L'échantillon est ainsi passé de 4800 familles en 1968 à plus de 9 000 en 2009, soit plus de 70 000 individus cette même année, balayant plus de 40 ans de leur vie.

Bien que ces deux enquêtes répondent à une méthodologie d'enquête proche, le PSID est a priori moins surdéterminé par l'administration fédérale que le recensement puisque ce panel est géré par une université. En revanche, en reprenant la variable « race » du recensement, les chercheurs ne peuvent la modifier.

Les acteurs de la production des données issues d'enquêtes

Ces deux enquêtes partagent un point commun majeur pour qui analyse les discriminations. La définition normative de la variable statistique « race », telle qu'elle est définie *ex-ante* par les questionnaires, y est identique : elle repose sur les catégories du recensement. Restituer l'histoire de cette catégorie permet de mesurer la variation et le poids des choix des concepteurs des questionnaires des enquêtes dans la production de cette donnée, et par suite, dans son utilisation par les chercheurs. Dans les deux bases, les chercheurs utilisent cette donnée comme variable et ne peuvent généralement pas en modifier la collecte, ou les contours conceptuels.

Les acteurs de la construction des données démographiques aux États-Unis

Aux États-Unis, la collecte des données démographiques concernant « l'origine » ou « l'identité » des individus a été tout d'abord le fait des administrations publiques. Étant donné l'origine et l'utilisation d'abord politique et administrative de ces données, ainsi que le nombre d'observations concernées, un ensemble important d'acteurs et d'institutions interagit pour les produire.

La question des données primaires – l'information collectée – est essentielle dans le cas des données démographiques puisque l'existence de ces catégories et leur très forte plasticité au cours des deux siècles écoulés sont le résultat de négociations politiques et administratives. Le recensement de la « race » est inscrit dans la Constitution américaine depuis 1790. Les données ethniques et raciales ont très largement évolué en fonction de la prépondérance de la « question noire », des changements de la politique d'immigration ou de changements plus spécifiques comme la possibilité de déclarer plusieurs « races », octroyée en 2000.

Une des premières caractéristiques des données qui nous intéressent ici est sa construction par les processus sociaux et politiques.

[Illustration 1]

Au cours de l'histoire, les catégories raciales ont mêlé l'ethnicité, la couleur de peau, la religion, l'origine nationale et la langue. Les contours de ces catégories et ce qu'elles désignent se sont très largement déplacés. De 1790 à 1850, seuls les « Blancs » et les « Noirs » étaient recensés, ces derniers étant divisés en « homme libre » et « esclave » ; les « indiens » (*natives*) ne seront comptabilisés pour la première fois qu'en 1890.

À partir de 1890, le recensement intègre dans sa structure la « loi de la goutte de sang ». Selon cette loi – « *one drop makes you whole* », – le fait d'avoir un ancêtre « noir » suffit à être classé dans la catégorie « noire », et ce, parfois, sans lien avec l'aspect visuel de la couleur de peau comme le montre l'illustration 1. Cette loi, jusqu'alors implicite, devient de plus en plus inapplicable à partir du moment où la généalogie d'une population de plus en plus grande et de plus en plus mobile devient impossible, rendant nécessaire la création de nouvelles catégories administratives. En effet, si la connaissance des ancêtres de chacun est possible dans de petites unités, elle ne l'est plus au niveau d'une nation entière – où la caractéristique visuelle l'emporte.

Or on voit bien comment cette divergence potentielle entre visuel et généalogie pose problème. Cette divergence brouille ce que W.E.B. Dubois appelle la « ligne de couleur » (*colorline*). Les enfants nés de relations entre maîtres et esclaves, bien qu'ayant un père blanc, ne peuvent être considérés *que* comme des « gens de couleur » afin que la hiérarchie sociale soit respectée. Cette tension s'explique aussi par le fait que l'octroi de droits supplémentaires pendant l'« ère de la reconstruction » (1865-1877) a conduit, paradoxalement, à la mise en place de la ségrégation institutionnelle¹¹, plus d'égalité menant paradoxalement à plus de séparation.

C'est au moment de ce brouillage des « lignes de couleur » que se développent des catégories plus fines dont l'objectif est de donner un contenu statistique au métissage : de 1850 à 1920, les recensements entendent ainsi dénombrer les « *quadroon* » (dérivé du français quarteron) et les « *octoroon* », individus ayant respectivement un grand parent ou un arrière grand parent africain américain (*cf.* Illustration 2).

Les distinctions ne tiennent qu'à la nécessité d'établir une hiérarchie, en haut de laquelle se situe la « race » blanche (*cf.* illustration 2). La nécessité de la distinction confrontée à l'ambiguïté visuelle de la « race » n'est pas

¹¹ L'arrêt *Plessy vs. Ferguson* de 1896 signe l'échec définitif des réformes d'après la guerre civile et la mise en place de la ségrégation institutionnelle selon la maxime « séparés mais égaux ».

nouvelle, comme le fait remarquer le compagnon de route de Toqueville (De Beaumont 1840, 12-13) et trouve des expressions actuelles dans des classements obéissants à d'autres critères. Ainsi, paradoxalement, Schor souligne que, à partir de 1990, la possibilité pour les enquêtés de cocher différentes « races », au lieu d'une catégorie « multi- raciale », s'est accompagnée, pour les personnes ayant coché « Blanc » ainsi qu'une autre race, de leur assignation systématique au groupe minoritaire des « autres races ». Le but était de parvenir à des effectifs nombreux et de pouvoir ainsi « produire les groupes de référence des politiques publiques » (Schor 2009 : 335-336). Il s'agit ainsi de l'« ultime avatar de la *one drop rule* [mais] à fronts renversés, puisque ce sont aujourd'hui les organisations représentant les minorités qui exigent que leur visibilité soit maximisée par le dispositif statistique » (*ibid.*).

[Illustration 2, 3 et 4]

Des catégories « raciales » à la variable « race »

L'assignation d'une « race » à un individu dépend de deux groupes de circonstances : d'une part de la définition des modalités possibles de cette catégorie « race », et, d'autre part, de la personne qui assigne cette modalité à l'individu concerné.

Le choix des différentes modalités possibles pour la « race » est central dans l'usage des données : au-delà de l'ordre hiérarchique et symbolique qui entoure une « identité », la forte variation de ce que l'on mesure – exemple de l'histoire mouvementée de la prise en compte du métissage – a des conséquences sur l'utilisation de la variable pour créer des indicateurs ou mesurer des phénomènes comme la discrimination.

Depuis 2010, il existe dans le recensement américain 15 catégories, avec la possibilité de cocher en plus « other Asian », « other Pacific islander » ou « some other race ». La catégorie « Hispanique, Latino ou d'origine espagnole » est introduite en 1970 et concerne tous les individus qui peuvent également déclarer une « race » (cf. illustration 4).

L'histoire de la catégorie raciale se reflète parfaitement dans le cas de la CPS et du PSID. La variable « race » dans le PSID a connu une nette évolution entre 1968 et 2015, parfaitement superposable aux évolutions des recensements décennaux. Les illustrations 8 et 9 mettent en lumière les changements normatifs qui ont eu lieu, avec un élargissement des modalités possibles de réponse pour la variable « race », parallèlement aux évolutions observées dans le recensement décennal. L'illustration 9 montre également comment plusieurs noms peuvent définir une même catégorie (voir catégorie 2 : « *Black, African-American, or Negro* »). En effet, selon les contextes et

cultures dans lesquels les individus ont évolué, leur identification à un groupe peut se faire au travers de noms d'appartenance différents.

[Illustrations 8 et 9]

Comme on l'a vu, la construction de ces modalités (ici les modalités possibles d'assignation d'une « race » à un individu) fait intervenir de très nombreux acteurs (personnels politiques, administrations, chercheurs, citoyens) et découle de processus complexes, politiques, juridiques ou liés à la mobilisation de groupes de pression¹². Sa collecte également.

Les acteurs de la collecte de ces données furent d'abord l'armée d'enquêteurs (souvent des femmes) envoyés dans l'ensemble du pays afin de remplir les questionnaires du recensement. Jusqu'aux années 1960, la catégorie « race » fait l'objet d'un recensement par une tierce personne, le plus souvent un enquêteur du Census Bureau ; à partir de 1980, le recensement de la « race » s'effectue par auto-identification. L'assignation d'une modalité à un individu, ici la « race », peut varier selon la personne qui procède à cette assignation. Le passage de l'assignation par une tierce personne à l'auto-déclaration met en évidence le rôle important de la perception de la « race » par les différents acteurs du processus¹³.

Les modalités et leur nombre ne sont pas sans impact sur la place laissée au chercheur dans la construction des variables, dès lors que l'on se place dans démarche hypothético-déductive telle que celle qui domine actuellement en économie. En effet, plus le nombre de modalités d'une variable est faible, plus il est probable que le découpage disponible soit incompatible avec le découpage que prescriraient les modèles du chercheur. A l'inverse, plus le nombre est élevé, plus l'éventail des regroupements possibles est large et plus le poids normatif de la théorie dans les données construites est important. C'est le cas par exemple pour l'étude des descendants d'immigrés qui requiert l'identification du pays d'origine de leurs parents ; or la CPS propose 161 modalités sur le pays d'origine. Une étude approfondie sur les descendants des 161 groupes d'immigrés ne

¹² Pour une histoire générale de ces négociations, voir Schor 2009, en particulier l'histoire de l'échec de la constitution d'une catégorie raciale « mexicaine-américaine » (Schor 2009 : 248).

¹³ Un certain nombre d'études mettent en évidence les différences de perception de la couleur de peau, notamment entre l'attribution faite par les enquêteurs et l'auto-identification des enquêtés, même lorsque que des échelles de couleur matérielles reproduisant les différentes nuances sont utilisées. Voir notamment Telles (2002).

pouvant être réalisée¹⁴, des regroupements d'origine doivent être faits et sont laissés à la discrétion des chercheurs.

Le rôle du chercheur au sein de la production des données : en aval

La production des données obéit aussi à des contraintes techniques. La mise en série de ces données du recensement met en jeu une diversité importante d'acteurs : une fois les questionnaires complétés, ils doivent être acheminés puis traités une première fois afin d'en extraire les informations pour énumération. Les données ont d'abord été enregistrées et codées par des opératrices de saisie, avant l'informatisation du processus. L'enregistrement et le codage se font à l'aide d'un support : il s'agit d'abord de tableaux manuscrits, puis de cartes perforées¹⁵, utilisées jusqu'aux années 1970, puis de bandes magnétiques, avant d'être informatisées¹⁶. À ces supports sont liés également des appareils de lecture – de la mécanographie à l'informatique – qui sont aussi parfois ceux utilisés pour effectuer des calculs. Les bases de données informatiques (le terme *database* apparaît en 1964) vont considérablement modifier les possibilités du calcul et donc, de la mesure à grande échelle ; essentiellement en multipliant le nombre de caractères statistiques et le nombre de modalités des variables, mais également la vitesse de calcul.

La conservation des données d'enquête pose la question de leur gestion. D'une part, les données d'enquête sur lesquelles reposent les mesures de la discrimination sont des enquêtes longitudinales. Cela implique de réutiliser et donc de conserver les questionnaires utilisés dans l'ensemble des collectes ainsi que les dictionnaires de variables, c'est-à-dire le document fournissant les informations relatives aux données. Par exemple, le dictionnaire des variables procure, pour chaque variable, le nom de la variable, son libellé, une liste des modalités de la variable, des renseignements sur l'origine de la variable, le nombre d'observations par variable, etc. De plus, la dimension longitudinale permet l'introduction au cours du temps de nouvelles variables,

¹⁴ Pour des raisons pratiques (contraintes de temps et d'espace) mais aussi d'effectifs et ainsi de significativité statistique (certains groupes rassemblent trop peu d'individus).

¹⁵ Les cartes perforées ont été inventées au 17^{ème} siècle pour les besoins de l'industrie textile. Développée par la suite pour améliorer l'efficacité du recensement américain – notamment par Herman Hollerith en 1884, qui l'utilisera pour le recensement de 1890 – cette technologie sera exploitée à la fois par l'État fédéral et par de grandes entreprises américaines, dont IBM (*International Business Machines Corporation*).

¹⁶ Les cartes perforées ont d'abord été exploitées *via* des tabulatrices (*cf.* Illustration 5), puis des machines électro-comptables. Ces technologies – la mécanographie – apparaissent au 19^{ème} siècle pour mourir dans les années 1970. Ce processus est bouleversé par l'invention des disques durs en 1956.

obéissant à de nouveaux objectifs de politique publique ainsi qu'à de nouvelles normes culturelles. Il n'y a pas de différence entre la CPS et le PSID quant à ces deux aspects (gestion et conservation des bases de données).

Les gestionnaires des données peuvent être des archivistes, des administrateurs ou des techniciens. L'entretien des données passe par la correction des erreurs, le stockage à proprement parler, mais concerne aussi la mise à disposition de ces bases de données. Renouvelées tous les dix ans¹⁷, les données de recensement sont essentiellement gérées de manière administrative. La mise à jour et la commensurabilité des données restent les enjeux principaux. Leur origine administrative détermine leur pérennité, en termes de suivi et de disponibilité pour les recherches futures.

Enfin, des bases de données aussi conséquentes posent la question de leur coût. Cette dernière question se pose différemment si le producteur est l'État (producteur considéré comme pérenne¹⁸) ou un acteur privé : les producteurs peuvent décider de cesser la production de données et donc empêcher le suivi des objets à mesurer.

La collecte et la mise en forme de la catégorie raciale aux États-Unis ont posé les bases du développement de l'étude des inégalités et des discriminations raciales¹⁹. Ce détour historique permet de mettre en évidence la toile de fond américaine, les chercheurs héritant du système administratif, des catégories légales et des techniques statistiques.

Une des conséquences moins visible de la gestion administrative des bases de données concerne la disponibilité de certaines variables. Par exemple, le PSID n'intègre pas de variables sur le secteur d'activité, qui représente pourtant un facteur explicatif majeur des différences de salaire aux États-Unis. La décision de l'administration de ne pas inclure certaines variables dans les enquêtes s'avère déterminante pour la modélisation théorique : l'absence de variables, ou le mode de mesure lui-même contraint la formulation du modèle théorique effectivement utilisable.

Les bases de données sont collectées pour servir à des fins administratives et de recherches (publiques, académiques ou privées) et leur existence s'inscrit dans la durée. Les concepteurs des enquêtes n'ont, en

¹⁷ La majorité des recensements sont décennaux mais peuvent aussi, plus rarement, être quinquennaux.

¹⁸ Ce producteur pérenne peut néanmoins décider, pour des raisons budgétaires, de couper les ressources allouées comme ce fût le cas en 1982 où 11 000 ménages ont été supprimés de la CPS, sur les 71 000 interviewés en 1981 (Bregger 1984).

¹⁹ Il nous semble que l'existence des statistiques raciales explique en partie le développement d'études quantitatives sur la discrimination en économie. À titre de comparaison, le terme de discrimination apparaît dans les débats vers la fin des années 1980 en France (Fassin 2002; 2006) alors qu'il se diffuse largement aux États-Unis dans les années 1950.

outre, pas forcément en tête les modèles théoriques ou les questions des chercheurs.

Produire des données expérimentales pour mesurer les discriminations

Deux exemples d'expérimentations pour mesurer les discriminations raciales

Les méthodes expérimentales appliquées à l'étude des discriminations se sont en partie développées pour pallier les insuffisances des analyses utilisant des données d'enquête. Ce n'est pas le cas de l'ensemble des expérimentations en économie (Cot et Ferey 2016). Dans le cas de l'étude des discriminations, la collecte de données expérimentales vise à produire des informations sur les *comportements* discriminatoires, par définition non-observables dans les données d'enquête.

Nous distinguons d'abord le processus de production des données issues d'expérience de terrain de celles issues d'expérience de laboratoire (section 2.1.) avant de préciser la fonction des différents acteurs dans la production de ces données, fonction clairement distincte de celle des acteurs produisant les données d'enquête (section 2.2.). Le principal résultat observé est une « fusion » des différents acteurs présentés précédemment dans le cas des données d'enquêtes en un seul acteur : l'équipe de recherche²⁰.

Mesurer les discriminations avec des expériences de terrain

L'expérience de terrain sur les discriminations raciales la plus citée a été menée par Bertrand et Mullainathan (2004). La méthode employée est celle du *testing* : elle consiste à envoyer des CV fictifs en réponse à des offres d'emploi réelles et à recueillir les réponses des recruteurs. Les auteurs cherchent à déterminer si les individus fictifs sont traités différemment en fonction de leur appartenance de groupe – l'expérience s'intéresse aux Africains-Américains et aux « WASP²¹ » aux États-Unis.

²⁰ Dans cette section, nous utilisons indifféremment les termes chercheurs, auteurs, expérimentateurs pour désigner l'ensemble des membres d'une l'équipe de recherche dont la taille peut varier.

²¹ L'acronyme WASP signifie *White Anglo-Saxon Protestant* et désigne généralement aux États-Unis les « blancs » américains d'origine anglaise et protestante, souvent rattachés par extension aux premiers colons du Nouveau Monde.

Les méthodes expérimentales en économie

Les expériences ont lieu « dans un environnement contrôlé afin de reproduire artificiellement une situation reflétant les conditions réelles, mais dans lequel le chercheur peut manipuler plusieurs paramètres » (Eber & Willinger, 2012 : p.4). Les participants à l'expérience sont souvent appelés *joueurs* ou *sujets*. À la différence des producteurs des données d'enquête, l'expérimentateur peut, grâce au cadre expérimental, manipuler *l'environnement* et *les règles institutionnelles*. L'environnement fait référence à certaines caractéristiques des joueurs (par exemple leur revenu au début de l'expérience appelé dotation initiale) et au cadre dans lequel les décisions sont prises (tous les individus prennent des décisions dans le même lieu, après avoir entendu les mêmes instructions, etc.). Les règles institutionnelles régissent les interactions entre les joueurs.

Pour mesurer les discriminations, les résultats d'expériences antérieures et des hypothèses théoriques permettent de décrire les résultats attendus d'un jeu en l'absence de discrimination. Les écarts par rapport à ces décisions observées en laboratoire sont alors associés à de la discrimination. C'est en contrôlant l'environnement et les règles institutionnelles que l'expérimentateur teste si une règle spécifique induit une décision déviante des prédictions théoriques.

Les données expérimentales peuvent être issues d'expérience en laboratoire ou sur le terrain. Les expériences en laboratoire rassemblent des participants dans une salle, où ils sont isolés les uns des autres et ils n'interagissent que *via* un réseau informatique (Eber & Willinger, 2012, *cf.* illustration 10). Le réseau informatique permet à l'expérimentateur de contrôler la communication entre les participants et de récolter les données issues des décisions des participants.

[Illustration 10]

Les expériences sur le terrain ont lieu dans un environnement réel dit « naturel ». Les participants issus de la population générale (qui ne savent pas toujours qu'ils participent à une expérience) prennent des décisions économiques réelles. Les expériences sur le terrain réduisent les artifices inhérents aux expériences en laboratoire, *i.e.* les biais de sélection des sujets qui ont des caractéristiques individuelles particulières (un sujet standard est un étudiant en économie ; il a du temps en pleine journée, par exemple) par rapport à la population générale, ou encore par rapport au réalisme du type de décisions prises en laboratoire (à quelle prise de décision en environnement naturel correspond le jeu du dictateur ?).

Pour générer une différence d'ethnicité perçue par l'employeur, des prénoms²² aux sonorités africaines-américaines ou « WASP » sont

²² Les noms de famille sont aussi manipulés pour correspondre à la « race » mais font l'objet d'une sélection moins rigoureuse.

aléatoirement associés à des CV en tous points similaires. Pour chaque annonce, quatre CV correspondant à la fiche de poste sont sélectionnés (deux moins qualifiés et deux plus qualifiés). Un prénom africain-américain est aléatoirement associé à un CV moins qualifié et à un CV plus qualifié. Un prénom WASP est associé aux deux autres CV. Le but est également de tester l'effet de l'ajout d'une information supplémentaire, signalant un plus haut niveau de qualification.

Dans cette expérience, les CV sont envoyés en réponse à des offres d'emploi réelles, publiées dans des journaux à Chicago et à Boston. Les expérimentateurs ont envoyé quatre réponses (CV) aux 1 300 annonces sélectionnées, soit 5 200 CV. Les employeurs potentiels peuvent répondre par téléphone ou par courriel, un numéro de téléphone et une adresse de courriel sont créés pour chaque candidat. Le contenu de la communication éventuelle entre employeur et candidat (absence de réponse, délai de réponse, proposition d'entretien, etc.) est ensuite enregistré par l'équipe de recherche.

Le protocole vise à garantir que les différences de taux de réponse des employeurs aux candidats fictifs soient attribuables uniquement à la différence de « race » perçue par les employeurs. Le chercheur interprète les différences systématiques de réactions des employeurs comme le résultat de discriminations. Les résultats indiquent que les candidats avec un prénom « WASP » ont 50 % de plus de chances de recevoir une réponse d'un employeur que ceux qui ont un prénom à consonance africaine-américaine, toutes choses égales par ailleurs. Le taux de réponse augmente avec la qualité de la candidature de manière plus importante pour ceux avec un prénom « WASP » que pour ceux qui ont un prénom africain-américain²³.

Cette méthode a été très largement appliquée en économie et en science politique à partir des années 1970 – pour une méta-analyse récente de plus de 28 expérimentations concernant la discrimination à l'embauche aux États-Unis, voir Quillian *et al.* (2017)²⁴.

Mesurer les discriminations dans un laboratoire

L'expérience en laboratoire permet d'observer des décisions effectuées dans un environnement contrôlé par l'expérimentateur. Les sujets de l'expérience interagissent en jouant les uns contre les autres. Leurs interactions sont régies par des règles du jeu qui indiquent notamment les

²³ Ce dernier élément révèle l'existence de discrimination de type statistique (Arrow 1973) avec des effets d'apprentissage plus importants pour les individus perçus comme « WASP ».

²⁴ La première occurrence semble être l'étude britannique (non randomisée) de Jowell et Prescott-Clarke (1970). Pour deux exemples récents concernant la discrimination à l'embauche en France, voir Edo et Jacquemet (2013) et Valfort (2017). Pour une perspective historique, voir Chassonnery-Zaïgouche (2014, chapitre 5).

gains auxquels ils peuvent prétendre. En général, les joueurs n'interagissent pas directement, tout se passe *via* l'interface de l'ordinateur, ils ne connaissent pas l'identité des autres sujets²⁵. Chaque joueur ne voit pas les autres joueurs prendre des décisions car les ordinateurs sont dans de petits habitacles. Le laboratoire n'a pas vocation à imiter ce qui se passe dans la vie courante, puisqu'il s'agit justement de confronter les sujets à des situations les plus épurées possible. Cependant, les décisions prises par les participants sont bien des décisions économiques réelles, puisque, de leur choix, dépendent des gains monétaires. Le principe sous-jacent est que les sujets vont chercher à maximiser leurs gains dans les jeux qui leur sont proposés. Notons qu'il n'est pas explicitement indiqué aux joueurs dans les instructions qu'ils doivent maximiser leurs gains, mais il est implicitement supposé qu'ils le feront.

C'est ce type d'expérience que mènent Fershtman et Gneezy (2001) pour mesurer la discrimination « ethnique » en Israël²⁶. Nous avons délibérément choisi l'exemple de cette expérience parce qu'elle nous éloigne du contexte américain, et qu'il permet de penser le rapport entre la catégorie raciale légale et la catégorie utilisée par le chercheur. Le bureau central des statistiques israéliennes classe la population en fonction de la religion, essentiellement en juifs et non-juifs²⁷. Les catégories ethniques mobilisées dans l'étude appartiennent à la même catégorie, i.e. juive : il s'agit des Ashkénazes (immigration américaine et européenne) et les juifs originaires d'Afrique du nord et du Moyen Orient, seconde catégorie qu'ils nomment « *eastern jews* ». Les auteurs ont choisi de rassembler dans une seule catégorie les juifs du Moyen Orient (« *eastern or oriental jews* » en anglais et « *mizrahim* » en hébreu) et les juifs d'Afrique du nord (séfarades). L'expérience a eu lieu dans trois laboratoires d'économie expérimentale universitaires en Israël. Les sujets sont 966 étudiants israéliens de différentes universités. L'origine ethnique est indiquée par le vrai nom de famille de chaque joueur. Le nom est indiqué dans les instructions de l'expérience, c'est-à-dire qu'il est connu du jour avant que les joueurs prennent des

²⁵ Les sujets peuvent également jouer contre un programme informatique.

²⁶ Tout comme la catégorie de « race », la notion d'« ethnique » est une notion construite socialement, souvent de manière particulière au contexte national. Nous utilisons ici le terme des auteurs de l'article. Pour une mise en perspective de l'usage, voir Safi (2013).

²⁷ Il existe en fait deux types de classement : par la religion et par le « groupe » (« *population group* »). Dans les statistiques par religion, les catégories sont : juifs, musulmans, chrétiens (qui comprend les chrétiens arabes et les chrétiens non-arabes), les Druzes et les personnes qui ne sont pas classées selon leur religion. Le classement par « groupe », le plus utilisé, procède de trois catégories : juifs, arabes, et « autres ». Jusqu'en 1995, les « autres » étaient inclus dans la catégorie « arabes ». Depuis 1995, la catégorie « autres » fait référence aux chrétiens non-arabes, aux membres d'autres religions et aux personnes qui ne sont pas classées selon leur religion. Voir le *Statistical Abstract of Israel*, CBS (2010 : 2, tabl. 2.3-2.9).

décisions. L'affectation des partenaires de jeu est aléatoire, mais entre deux joueurs n'appartenant pas à la même université.

L'expérience est constituée d'une succession de jeux (jeu de la confiance, jeu de l'ultimatum, jeu du dictateur) durant lesquels les sujets prennent des décisions stratégiques en fonction de ce que font les autres sujets. Les différences de réaction observées entre des joueurs appartenant à des groupes différents sont interprétés comme différents degrés de générosité ou de défiance ; si les biais contre un groupe sont systématiques, les auteurs parlent alors de discrimination. La différence de montant monétaire échangé mesure l'intensité de cette discrimination.

Une rapide description du jeu de la confiance, premier jeu utilisé dans cette expérience, permettra de mieux comprendre. Ce jeu implique deux joueurs (A et B) qui jouent successivement un coup. L'expérimentateur confie une dotation initiale (un certain montant d'argent) au joueur A. Le joueur A peut donner une partie de sa dotation au joueur B, dans la proportion qu'il souhaite, ou garder l'ensemble de la dotation. Si le joueur A décide de transférer une partie de sa dotation initiale au joueur B, celui-ci reçoit un supplément de dotation : il dispose, pour jouer au coup suivant, du triple du montant transféré par le joueur A. C'est ensuite au joueur B de jouer : il peut décider, à son tour, de donner au joueur A une part de ce montant augmenté ou choisir de garder la somme pour lui. Avec de telles règles, les gains des joueurs sont d'autant plus élevés que les deux joueurs coopèrent : le montant en jeu au 2^{ème} coup est maximum si le joueur A transfère la totalité de sa dotation initiale au joueur B, mais le premier risque de tout perdre si le second ne coopère pas en retour. Le jeu peut s'arrêter là, ou être reproduit en demandant aux sujets d'échanger de rôle au sein de la même session expérimentale. Les joueurs ont connaissance du déroulement et des règles du jeu, et ils peuvent en déduire que leurs gains seront supérieurs s'ils coopèrent. Le montant que le joueur A transfère au joueur B est un indicateur de la confiance que ce joueur accorde à l'autre joueur et indique leur capacité à coopérer. L'objectif des chercheurs est de déterminer si cette volonté de réciprocité (partage des gains) varie en fonction de l'origine des joueurs qui se font face.

Les résultats de cette expérience, que les auteurs ont diffusés auprès des étudiants dans un but pédagogique (Fershtman et Gneezy, 2001 : 374), sont sans appel : le transfert d'argent moyen des Ashkénazes vers les juifs de l'est est moitié moins important que le montant transféré aux hommes Ashkénazes²⁸.

²⁸ Les auteurs n'ont trouvé aucune discrimination de la part des femmes. L'effet du nom non-Ashkénaze est moins important pour les femmes dans leur interaction avec les hommes.

Les éléments principaux de ce protocole expérimental ont été réutilisés récemment pour mesurer les discriminations raciales et envers les musulmans en France (Adida *et al.* 2016).²⁹

Les acteurs de la production des données issues d'expérimentations

Dans les deux types d'expériences présentées, le rôle de l'équipe de recherche est prédominant dans la production des données à travers la conception du protocole expérimental (section 2.2.1). Au sein de ce protocole, la variable « race » n'est plus héritée d'un appareil statistique administratif mais constitue un des objets de la recherche à part entière, dont les modalités peuvent être décomposées par les chercheurs eux-mêmes (section 2.2.2). A la multitude d'acteurs de la production et de l'entretien des bases de données mentionnés pour les données d'enquête, s'oppose l'intervention d'équipes de recherche restreintes pour les données expérimentales (section 2.2.3).

Produire des données à partir d'un protocole expérimental

Dans le cas des données expérimentales, les chercheurs construisent un protocole pour produire leurs propres données. Le protocole expérimental indique précisément la méthode employée et le rôle de chacun. Il définit la durée, le lieu et le type de sujets et de population concernés, le type d'environnement et le type d'interactions sociales observées. Les chercheurs doivent justifier qu'aucune des décisions relatives au protocole ne biaise les résultats attendus de l'étude. Une fois le protocole défini, celui-ci est mis en œuvre par l'ensemble de l'équipe de recherche.

Dans le cas du testing, il s'agit de créer les différents CV utilisés pour les types d'emplois sélectionnés (vendeur, personnel administratif, employés de bureau et dans des services clients), mais également de mener l'expérience à proprement parler (répondre aux annonces et enregistrer les réponses). À l'issue de l'expérience, une réponse de l'employeur est associée à chaque CV envoyé.

Dès lors, l'employeur (non fictif) peut être considéré comme un acteur à part entière dans le processus de production des données. Bien que celui-ci ne soit pas au courant qu'il participe à une expérience, c'est bien de sa réaction que va dépendre l'information associée au CV. Cette information peut-être soit un appel ou un email proposant un entretien ou un rejet de la

²⁹ Les auteurs de cette étude ont fait à la fois une expérience de terrain et une expérience de laboratoire.

candidature (ou une absence de réponse par exemple). L'article ne développe pas la manière dont cette donnée est inscrite dans la base de données, mais on peut proposer une représentation schématique de la base de données constituée qui permet d'enregistrer la réponse de l'employeur pour chaque annonce et pour chaque CV envoyé (tableau 1).

Employeur	CV envoyé	Réponse employeur	Date d'envoi du CV	Ville
Entreprise A	Qualifié + WASP	Entretien	1/10/2001	Boston
Entreprise A	Non Qualifié+ WASP	Entretien	1/10/2001	Boston
Entreprise A	Non Qualifié + Africain Américain	Rien	1/10/2001	Boston
Entreprise A	Qualifié + Africain Américain	Entretien	1/10/2001	Boston
Entreprise B	Non Qualifié + WASP	Négative	1/04/2002	Chicago
Entreprise B	Non Qualifié + Africain Américain	Négative	1/04/2002	Chicago
Entreprise B	Qualifié + Africain Américain	Négative	1/04/2002	Chicago
Entreprise B	Qualifié + WASP	Entretien	1/04/2002	Chicago

Tableau 1 Données primaires fictives illustrant la collecte des données de l'expérience

On voit apparaître (schématiquement) dans le tableau 1 une base de données où la première ligne indique le nom de l'information donnée dans chaque colonne. Les autres lignes correspondent chacune à un CV envoyé. Il y a 4 CV envoyés à chaque entreprise en réponse à une annonce. Chaque CV correspond à la combinaison d'un niveau de qualification et d'une catégorie raciale. On observe, pour chaque CV, l'entreprise à laquelle le CV a été envoyé, la réponse de l'employeur, la date d'envoi du CV, la ville dans laquelle l'annonce a été publiée. Les données primaires sont ainsi mises en séries par l'équipe de recherche.

Pour commencer l'analyse il est nécessaire de mettre en forme les données primaires afin qu'elles puissent être lues par le logiciel, la plupart des logiciels de traitement statistique ne pouvant traiter que les chiffres. Il faut donc coder chaque type de réponse possible. Le passage de la base de données primaires à une base de données analysable est souvent appelé le « nettoyage ». Le nettoyage est une succession d'actions qui consistent à

créer de nouvelles variables à partir des informations présentes et à recréer des catégories dans les variables existantes. L'enregistrement des données, leur codage et le « nettoyage » sont effectués par l'équipe de recherche.

Dans le cas des expériences en laboratoire, le protocole indique l'ensemble des étapes de l'expérience depuis l'arrivée des sujets jusqu'à leur départ (dans certains protocoles, les sujets ne doivent pas se voir, ou se parler).

Contrairement aux expériences de terrain, il n'est pas rare de mener une expérience de laboratoire pilote afin de tester un protocole. Il s'agit de tester l'expérience afin d'observer si l'expérience programmée correspond à ce qui est indiqué dans le protocole, de déceler les erreurs et de vérifier que les données primaires récoltées permettront de produire l'analyse escomptée.

Les sujets sont généralement rémunérés pour leur participation (le montant est par exemple de 5 euros au Laboratoire d'Economie Expérimentale de Paris 1) ainsi qu'en fonction des décisions prises pendant l'expérimentation (gratification variable déterminée par la performance ou les décisions). Les auteurs doivent décider des montants de gratifications possibles à chaque jeu et des conditions sous lesquelles les sujets obtiennent les gains possibles. Les expérimentateurs choisissent les visuels et le nombre d'informations sur chaque écran, la possibilité pour un joueur de revenir sur ses décisions, l'ordre des jeux (randomisé ou selon un ordre précis), etc. Une attention particulière est apportée aux instructions communiquées aux sujets. Ces instructions sont généralement disponibles sous une forme imprimée ou sur l'ordinateur mais sont généralement lues et expliquées par l'expérimentateur. Les instructions sont toujours publiées avec l'article.

À la différence des expériences de terrains, la collecte des données et leur codage sont généralement automatisés à travers l'usage d'un logiciel dédié (*z-Tree* ou *ComLabGames* par exemple). Le programmeur, en général salarié du laboratoire, est responsable de la traduction du jeu théorique en séquences de jeu informatique. Ces séquences de jeux produisent les données directement enregistrées par le logiciel, qui les met en série (et la plupart génère une synthèse sous forme de tableau).

Contrairement au chercheur travaillant sur des données d'enquête, les chercheurs procédant à des expériences sont initiateurs de la récolte des données et des catégories de données primaires.

Numéro de la session	Identifiant du Joueur	Offre du joueur A	Revenu mensuel	Age
1	1	10	1000	25
1	2	5	2000	32
1	3	6	3000	18
1	4	3	1000	21

Tableau 2 Données primaires fictives illustrant les décisions et les caractéristiques sociodémographiques de plusieurs sujets

Construire la variable « race »

La longue histoire de la variable « race » aux États-Unis est ici à mettre en relation avec des expérimentations dont l'usage d'une catégorie raciale ou ethnique se fait en relation avec la création d'un « signal ». Qu'il s'agisse des CV fictifs ou des noms des sujets dans l'expérience de laboratoire, l'information diffusée ne concerne pas la catégorie du recensement à laquelle appartient les individus mais renvoie à des catégories utilisées par les acteurs pour en déduire une affiliation de groupe : dans les deux cas ici, il s'agit d'un nom. Dans l'expérience de terrain, les noms renvoient à un groupe qui a une existence juridique à travers le recensement : les Africains Américains aux États-Unis. Dans l'expérience de laboratoire, les noms renvoient à un groupe construit par les chercheurs. La référence à ce groupe existe dans la société israélienne et préside à sa structuration hiérarchique. Il ne correspond pas à une catégorie statistique officielle mais à plusieurs sous-catégories. Les auteurs ont rassemblé dans une même catégorie des juifs de différentes origines qui partagent la caractéristique *de ne pas être d'origine Ashkénazes*, soit de ne pas appartenir au groupe jugé dominant dans la société israélienne. Le protocole expérimental permet ainsi d'analyser l'usage d'une catégorie visiblement utilisée par les acteurs, en mettant en évidence le fait que des signaux tel que le nom de famille ont un impact sur leur prise de décision en situation stratégique.

L'assignation ou l'auto-assignation à une catégorie raciale comme base de la récolte d'une donnée sur la « race », ensuite intégrée sous forme de variable dans une analyse économétrique est remplacée dans le cadre expérimental par la création d'un signal. Les patronymes mais également des photographies sont régulièrement utilisés. Des éléments contextuels sont

également utilisés pour décomposer les éventuelles significations de l'affiliation à un groupe : ajout de la religion comme pratique régulière ou non, adhésion à une association communautaire, informations suggérant les différentes langues parlées et le degré d'aisance dans chacune d'elle. Ces informations sont signalées dans le corps du CV s'il s'agit d'un *testing*, sur des profils d'utilisateurs ou de joueurs dans le cas d'expériences de laboratoire. L'intérêt est de pouvoir tester l'effet de différents aspects de l'affiliation de groupe. Par exemple, l'ajout de compétences linguistiques spécifiques annule parfois les comportements discriminatoires envers des individus aux patronymes perçus comme étrangers (Edo et Jacquemet 2013) mais pas systématiquement.

La construction des modalités de la variable d'affiliation de groupe, si elle obéit à des contraintes techniques et contextuelles (les identités signalées doivent être crédibles), dépend entièrement de l'équipe de recherche. En général, la constitution de la variable raciale dans les expérimentations se fonde sur des enquêtes préliminaires à l'expérience elle-même. Ces enquêtes *ex-ante* sont effectuées auprès d'un échantillon représentatif de la population générale pour savoir si les prénoms choisis sont associés à un certain groupe. Dans le cas des expériences de terrain, aucune mesure du degré d'association du nom à une « race » n'est effectuée *ex post*. C'est parfois le cas dans les expériences de laboratoire.

À titre d'exemple, dans le cas de l'enquête de Bertrand et Mullainathan, le choix des prénoms s'est fondé sur l'analyse de la fréquence à laquelle un prénom a été donné entre 1974 et 1979 dans l'un et l'autre des groupes raciaux. Pour vérifier la pertinence de leur choix, les chercheurs ont effectué une enquête à Chicago : il a été demandé à 30 personnes dans des espaces publics d'associer des caractéristiques (dont la race) à un prénom. Les catégories raciales proposées étaient Blanc, Africain Américain, Autre ou Ne peut pas dire. La plupart des personnes interrogées ont attribué la race attendue à chaque prénom. Les prénoms pour lesquels ce n'était pas le cas n'ont pas été utilisés.

La prédominance du rôle du chercheur dans la production des données expérimentales

L'article de Bertrand et Mullainathan mentionne 12 assistants de recherche. La division du travail n'est pas explicitée mais seuls deux auteurs signent la publication. Aucune précision n'est faite concernant l'expérience de Fershtman et Gneezy. La thèse de la « fusion » des multiples acteurs de la production des données dans le cas des expérimentations visait à accentuer une des différences saillantes vis-à-vis de la production des données d'enquête. Les types de bases constituées sont de natures différentes.

Les acteurs institutionnels engagés dans la production de base de données d'enquête représentative sont en charge de la pérennité de ces bases. Les bases de données constituées lors d'expériences se composent de l'ensemble des décisions prises dans chaque jeu, par chaque sujet, dans un contexte spécifique. Contrairement aux données d'enquêtes observationnelles, elles sont, en général, créées pour un usage unique. Ces bases de données ne sont pas publiques et ne sont donc pas, en théorie, accessibles à d'autres chercheurs qui voudraient les utiliser.

La transmission des données peut aussi répondre à la volonté de mettre en série plusieurs expériences, mais cette mise en série n'a de sens que dans le cas d'une extension géographique et temporelle. Le but de ces extensions est de s'interroger sur l'existence d'un tel phénomène dans d'autres zones géographiques ou pour d'autres populations (population générale, étudiants, etc.), ou encore, sur la persistance de ce phénomène dans le temps. S'ajoute à cela, le fait que pour prolonger l'analyse d'un comportement, il convient parfois de reproduire les résultats déjà obtenus par d'autres avant de tester l'effet d'un nouveau traitement. La déontologie scientifique veut donc que le protocole soit toujours précisé dans l'article publié et que les instructions fournies aux sujets soient disponibles dans l'annexe. C'est en effet la condition pour que l'étude puisse être reproduite : si un autre protocole et d'autres consignes sont utilisés, les résultats ne sont ni cumulables, ni comparables, car d'éventuelles différences de résultats pourraient être attribuées aussi bien à une évolution de la discrimination qu'à une différence de mesure liée au protocole.

Ainsi, dans le cas des enquêtes observationnelles, une administration, un centre de recherche ou une université sont en charge de produire, de conserver et de mettre les données ainsi que tous les documents nécessaires à leur compréhension, à la disposition des multiples utilisateurs potentiels. Dans le cas des données expérimentales, les données sont conservées par l'équipe de recherche qui les a produites et qui décide elle-même de les transmettre à d'autres chercheurs ou à un journal pour publier son travail. En pratique, il est parfois possible de demander aux auteurs de consulter leurs données. Certains les mettent même à disposition sur leur site internet. Certains journaux souhaitent y accéder avant publication et les diffusent si l'article est effectivement publié (en général sur leur site internet). Ce qui est mis à disposition d'autres utilisateurs potentiels, c'est le protocole expérimental. Élaboré spécifiquement pour répondre à une question de recherche particulière, il n'a cependant vocation à être utilisé que par un nombre restreint de chercheurs, pour produire des données cumulables. Une même équipe de recherche est donc, le plus souvent, le seul acteur de la production, de la conservation et de l'analyse des données.

Les données expérimentales, qu'elles soient produites sur le terrain ou en laboratoire, sont issues de l'observation de décisions provoquées par les

chercheurs eux-mêmes. Ainsi, la production des données, les opérations de transformation de ces informations ou des données primaires en variables, sont effectuées par les chercheurs ayant construit le protocole expérimental. L'équipe de recherche se trouve en aval et en amont de la production des données, de la conception du mode de collecte au codage, stockage et conservation des données. Contrairement à la répartition du rôle des acteurs dans le cas des données d'enquêtes, il y a, dans le cas des expérimentations, fusion entre ces différents acteurs.

* *

*

L'une des caractéristiques des données d'enquête est l'implication d'un nombre important d'acteurs dans leur production : des enquêteurs qui collectent les données primaires à ceux qui les enregistrent et les codent, des programmeurs (mécánographes, informaticiens) aux chercheurs (statisticiens, économistes, économètres) qui mettent en forme les données, en passant par les « usagers » qui les utilisent pour mesurer un phénomène ou évaluer une politique publique. Ces derniers sont rarement les commanditaires de la collecte. Par contraste, l'équipe de recherche est au centre de la production de données expérimentales. Ainsi, dans le cas des expériences de laboratoire ou de terrain, le nombre d'acteurs se réduit en général à une équipe de chercheurs, souvent les auteurs et des assistants, qui collectent les données primaires, fabriquent et gèrent la base de données et en infèrent une analyse.

La fusion entre tous les acteurs, de la production à l'analyse des données, change sensiblement les rapports entre théorie et empirie. Les données d'enquête répondent à une demande plutôt sociale et politique et leur mise en œuvre, à grande échelle, vise à donner des informations larges sur l'état d'un pays ou d'une région. Elles sont utilisées par les chercheurs, en aval, pour tester ou quantifier des analyses théoriques, mais, n'ayant pas été conçues dans ce but, elles ne comportent pas toutes les variables nécessaires à l'examen précis de certains modèles théoriques : l'utilisation de proxy, c'est-à-dire l'approximation d'une variable absente de la base de données par une autre variable statistique disponible, est fréquente. A l'inverse, les données expérimentales servent plutôt à tester un modèle théorique particulier, formulé en amont de l'élaboration des données. La production de données expérimentales permet aux chercheurs d'intervenir en amont et de disposer de données adaptées mais dont l'usage est le plus souvent unique.

Pour l'analyse des comportements de discrimination ethnique ou raciale, cette différence est de taille. Ainsi, l'enregistrement et la construction de la variable « race » dans le PSID suit les évolutions du recensement des États-

Unis, sans que les chercheurs ne puissent modifier sa définition, si ce n'est à la marge. À l'inverse, dans le cadre d'expériences, le chercheur n'est pas limité à l'utilisation de données raciales issues du recensement, qui recoupent des catégories juridiques, administratives ou statistiques, choisies par les producteurs du recensement en fonction des objectifs des pouvoirs publics. La variable raciale peut donc être construite en fonction d'autres marques d'identité perçues ou d'autres ressorts de la discrimination. C'est le cas dans les expériences décrites dans ce chapitre, qui utilisent une catégorie construite sur le signal que renvoient les prénoms, ou qui distinguent des groupes, qui ne correspondent pas à une catégorie administrative claire, mais correspondent à la réalité des préjugés à l'œuvre dans une société donnée.

Dès lors, il apparaît clair que le type de données disponibles influence largement les possibilités d'analyse et de calcul, ainsi que le type de phénomène que l'on peut mesurer. L'exemple de la mesure des discriminations raciales montre bien le passage de l'utilisation par le chercheur d'une catégorie raciale administrative à la construction d'une catégorie raciale expérimentale, déterminée par d'autres contraintes. Dans l'ensemble, la division du travail observée dans l'élaboration des données d'enquête tend à réduire la possibilité de manipulation des données par le chercheur, bien plus important dans l'élaboration des données expérimentales.

Bibliographie

- Adida, C. L., D. D. Laitin et M-A. Valfort (2016). « One Muslim is enough ! » Evidence from a field experiment in France. *Annals of Economics and Statistics*, 121-122 : 121-160.
- Armatte M. (1995). *Histoire du modèle linéaire : formes et usages en statistique et économétrie jusqu'en 1945*. Thèse de Doctorat, EHESS.
- Armatte M. (2010). *La science économique comme ingénierie : quantification et modélisation économique*. Presses des MINES, Paris.
- Arrow K. J. (1973). The Theory of Discrimination. In Ashenfelter, O. C. et Rees, A. (dir.), *Discrimination in Labor Markets*. Princeton, NJ : Princeton University Press.
- Bertrand M. et S. Mullainathan (2004). Are Emily and Greg More Employable Than Lakisha and Jamal ? *The American Economic Review*, 94(4): 991–1013.
- Blinder A. S. (1973). Wage Discrimination: Reduced Form and Structural Estimates, *The Journal of Human Resources*, 8(4): 436–455.
- Bregger J. E. (1984). The Current Population Survey: a historical perspective and BLS' role, *Monthly Labor Review*, 07(6): 8–14

- Central Bureau of Statistics (2010). Statistical Abstract of Israel 2010. http://www.cbs.gov.il/www/shnaton61/st_eng02.pdf [consulté le 27 Septembre 2017].
- Chassonnery-Zaïgouche, C. (2014). *Expliquer, mesurer, prouver : une histoire de l'économie des discriminations*. Thèse de doctorat. Université Paris 1 Panthéon-Sorbonne.
- Cot A. et S. Ferey (2016). La construction de « faits » économiques d'un nouveau type : éléments d'histoire de l'économie expérimentale. *L'actualité économique*, 92(1-2) : 1-31.
- De Beaumont G. (1840). *Marie ou l'esclavage aux États-Unis : tableau de mœurs américaines*. C. Gosselin, Paris.
- Eber N., et M. Willinger (2012). *L'économie expérimentale*. Paris : La Découverte.
- Edo A. et Jacquemet N. (2013). Discrimination à l'embauche selon l'origine et le genre : défiance indifférenciée ou ciblée sur certains groupes ? *Économie et Statistique*, 464-465-466 :155-172.
- Fershtman C. et U. Gneezy (2001). Discrimination in a Segmented Society: An Experimental Approach. *The Quarterly Journal of Economics*, 116(1): 351–377.
- Jowell R. et Prescott-Clarke P. (1970). Racial Discrimination and White collar Workers in Britain. *Race & Class*, 11 (4) : 397-417.
- Oaxaca R. L., (1973). Male-Female Wage Differentials in Urban Labour Markets, *International Economic Review*, XX, 14, 693–709.
- O'Connor, A. 2001. *Poverty knowledge : Social science. Social policy, and the poor in twentieth-century U.S. history*. Princeton : Princeton University Press.
- Quillian, L., Hexel, O., Pager, D., & Midtbøen, A. H. (2017). "Meta-Analysis of Field Experiments Shows No Change in Racial Discrimination in Hiring over Time." *Proceedings of the National Academy of the Sciences in the United States*. doi:10.1073/pnas.1706255114
- Safi, Mirna. 2013. *Les inégalités ethno-raciales*. La Découverte, Paris.
- Schor P. (2009). *Compter et classer. Histoire des recensements américains*. Paris: Editions de l'Ecole des Hautes Etudes en Sciences Sociales.
- Telles E. E. (2002). Racial ambiguity among the Brazilian population, *Ethnic and Racial Studies*, 25(3): 415–441.
- U.S. Bureau of Labor Statistics et U.S. Census Bureau (2006). History of the Current Population Survey, in *Design and Methodology, Current Population Survey Technical Paper 66*, 1–7.
- Valfort, M.-A. 2017. « La religion, facteur de discrimination à l'embauche

en France ? », *Revue économique*, 5 : 895-907.