

Comment expliquer la différence entre le score des blancs et des noirs américains aux tests d'aptitudes scolaires ?

Aux Etats-Unis, un écart important existe entre le score des blancs et le score des noirs aux tests d'aptitudes scolaires (SAT : Scholastic Aptitude Test). En ce qui concerne la lecture, un étudiant blanc américain a en moyenne un score de 527 et un étudiant noir américain un score de 428. Et idem pour les mathématiques (536 pour un étudiant blanc / 428 pour un étudiant noir) et l'écriture (515 / 417) (source : "National Center for Education Statistics").

Comprendre l'origine de cet écart est d'une importance cruciale pour réduire les inégalités raciales, très présentes aux USA. De nombreux économistes et chercheurs se sont donc intéressés à cette problématique du "Black-White Test Score Gap" et différentes hypothèses ont été avancées : (1) différence en ce qui concerne le niveau de pauvreté, (2) différence dans la structure familiale, (3) différence dans la qualité des écoles, (4) différence de traitement de la part des professeurs ou encore (5) différence purement génétique... Le *Captain'* met à l'honneur cette semaine les travaux de Roland G. Fryer Jr, économiste américain spécialiste des questions d'inégalités raciales et lauréat de la Médaille John Bates Clark 2015 (une sorte de mini Prix Nobel pour les économistes de moins de 40 ans).

Afin d'identifier les paramètres pouvant expliquer cet écart, il est nécessaire d'obtenir les données relatives aux résultats du test d'aptitude "SAT" (variable dépendante) et de croiser cela avec les données d'un grand nombre de variables explicatives (ethnicité, salaire des parents, école suivie...) sur un échantillon aléatoire d'une taille relativement importante. Prenons un exemple simple, avec donc notre variable dépendante (score au test SAT) et deux variables explicatives : l'ethnicité (noir ou blanc - le *Captain'* préférant le terme ethnicité à une traduction de l'anglais "race" pour éviter toute connotation) et le salaire des parents. Sur un échantillon de 2000 enfants, 1000 noirs et 1000 blancs, vous identifiez alors que le score moyen au test est de 500 pour les enfants blancs et de 400 pour les enfants noirs. Notre cher Jean-Marie pourrait alors s'empresser de conclure à partir de cela qu'il existe bien une "supériorité génétique de la race blanche". Mais si en analysant de plus près vos données (= avec un modèle économétrique tout simple), vous vous rendez compte, que, à salaire équivalent des parents, les noirs et les blancs ont un score équivalent au test, mais qu'en moyenne les parents noirs de l'échantillon ont un salaire inférieur aux parents blancs, la conclusion est toute autre. L'écart au test ne vient alors absolument pas d'une quelconque "supériorité génétique blanche", mais uniquement d'un écart de salaire entre les blancs et les noirs (l'écart de salaire pouvant impliquer un lycée de moins bonne qualité ou bien l'impossibilité pour les parents de payer des cours particuliers en cas de difficultés).

Dans l'article "Understanding the Black-White Test Score Gap in the First Two Years of School" (2004), Roland Fryer Steven Levitt (lauréat lui aussi de la médaille John Bates Clark) ont essayé d'identifier les variables expliquant ce "Black-White Gap", en analysant un échantillon de 20.000 enfants âgés entre 5 et 6 ans (kindergarten). En contrôlant par de nombreux facteurs sociaux-économiques, ainsi qu'en prenant en compte un grand nombre de variables de contrôle comme l'âge des parents, le nombre de livres à la maison, le poids à la naissance, la participation au programme de nutrition WIC, le nombre de fessées données par les parents (véridique), le niveau de sécurité du quartier d'habitation (et des centaines d'autres), les auteurs ont montré que l'écart entre les enfants noirs et les enfants blancs se refermait, c'est à dire que l'ethnie n'a pas d'impact causal sur le score au test SAT. Le tableau ci-dessous résume tout cela. Dans le premier modèle (1),

s'intéressant uniquement au lien entre le score au test SAT et l'ethnicité, le coefficient "Black" est significatif, négatif et élevé (-0,638, en jaune dans le tableau) montrant une nette différence entre le score des blancs et celui des noirs en mathématiques. Par contre, dans le modèle (5) incluant un grand nombre de variables de contrôle, le coefficient de "Black" a été divisé par 6 pour passer à -0,102 (forte réduction du "gap"). Idem en ce qui concerne le score au test de lecture (Reading) : dans le modèle (6) sans aucune variable de contrôle, le coefficient de "Black" est nettement négatif (-0,401) tandis qu'au contraire, en ajoutant toutes les variables de contrôle, le coefficient devient même positif, indiquant donc un score "toutes choses égales par ailleurs" meilleur pour les étudiants noirs que pour les étudiants blancs.

"The results of the previous section demonstrate that although black test scores lag whites by a large margin, the inclusion of a small number of covariates eliminates any systematic differences in the math and reading performance of whites and blacks entering kindergarten." Fryer Levitt

Cependant, dans les années suivant le jardin d'enfant, l'écart réapparaît très nettement ! Une hypothèse possible serait alors qu'il existe une forte différence dans la qualité des écoles suivies par les enfants noirs et par les enfants blancs. La ségrégation dans les écoles aux USA est en effet réelle : dans les 1000 écoles étudiées par Fryer et Levitt, 350 n'accueillent par exemple aucun enfant noir ! En réalisant différents tests économétriques en introduisant ce que l'on appelle un "fixed effect" pour identifier les variations à l'intérieur d'une même école accueillant à la fois des enfants blancs et des enfants noirs, les auteurs ont montré que la qualité de l'école expliquait une partie importante de l'écart aux tests d'aptitude au cours du temps (bien qu'une partie de l'écart reste inexplicée et que davantage de preuves empiriques seraient nécessaires).

Les statistiques ethniques, avec un peu d'économétrie, peuvent donc être utilisées pour combattre certains préjugés comme par exemple dans notre cas celui de la "supériorité génétique de la race blanche". En France, les statistiques ethniques sont interdites : impossible donc de réaliser une étude comme celle de Fryer Levitt. Bien que le débat sur la légalisation des "données raciales" soit loin d'être simple (voir par exemple "Le retour du débat sur les statistiques ethniques"), la collecte de statistiques ethniques, sans faire n'importe quoi, pourrait permettre de mieux comprendre certains problèmes, comme l'insécurité ou les discriminations à l'embauche, afin de mieux casser par la suite certains préjugés. Ou bien si finalement un préjugé est en fait réel, même en prenant en compte les facteurs socio-économiques et de nombreuses variables de contrôle, les statistiques ethniques pourraient permettre d'identifier les actions à mener, afin de sortir d'un débat purement idéologique en apportant alors des preuves empiriques.

"Je suis très attaché au modèle républicain. Et le refus de principe de l'Etat de prendre en compte ces différences part de bonnes intentions puisqu'il vise précisément à ne pas appliquer de traitement discriminatoire à telle ou telle catégorie de la population. Mais cette cécité volontaire devient dangereuse. [...] Ce combat contre la discrimination exige de prendre en compte ce critère racial. Non pas pour instaurer une discrimination positive au sens d'établir des quotas de recrutement dans l'administration, l'université, les entreprises, etc. Mais par exemple pour établir des statistiques ethniques qui permettent de mesurer l'ampleur de la discrimination négative. Et donc de la combattre résolument." - Robert Castel, sociologue

Conclusion : Le racisme, c'est mal ! Captain' Obvious