



ÉCLAIRAGES
ÉCIVICES

LES DISCRIMINATIONS EN RAISON DU GENRE ET DE L'ORIGINE SUPPOSÉE SUR DEUX PLATEFORMES COLLABORATIVES

G. CHAPELLE, P. DESCHAMPS, D. GLOVER, X. LAMBIN,
M. LAOUÉNAN, M. SESHIE, P. GRISOLIA, S. ALAYE, C. HENRY

MARS 2023

Les discriminations en raison du genre et de l'origine supposée
sur deux plateformes collaboratives

M A R S 2 0 2 3

Le Défenseur des droits soutient des activités d'études et de recherche afin de nourrir la réflexion et le débat public dans ses domaines de compétence.

Cette publication constitue une synthèse de la recherche intitulée « *Les discriminations en raison du genre et de l'origine supposée sur deux plateformes collaboratives : rapport d'étude* » et réalisée par une équipe de chercheurs du Laboratoire Interdisciplinaire d'Évaluation des Politiques Publiques (LIEPP) : G. Chapelle, P. Deschamps, D. Glover, X. Lambin, M. Laouéan, M. Seshie, P. Grisolia, S. Alaye et C. Henry.

Les opinions mentionnées dans cette publication n'engagent que ses auteurs et ne reflètent pas nécessairement la position du Défenseur des droits.

SciencesPo
LABORATOIRE INTERDISCIPLINAIRE
D'ÉVALUATION DES POLITIQUES PUBLIQUES

Avec le soutien de :

D
Défenseur des droits
RÉPUBLIQUE FRANÇAISE

RÉSUMÉ

Les plateformes numériques représentent une part croissante des échanges de biens et services en France. Jouant un rôle d'intermédiation entre le client et le prestataire, ces plateformes ont une intervention minimale : l'utilisateur sélectionne lui-même l'acheteur retenu parmi les réponses recueillies par son annonce, les vendeurs fixent eux-mêmes les prix des prestations et biens.

En modifiant les services proposés et les modalités de contact entre utilisateurs, ces nouvelles plateformes peuvent générer de nouvelles discriminations, mais pourraient aussi les limiter, notamment à travers la mise en place de procédures telles que la pseudonymisation ou la modification de leurs algorithmes.

Cette étude mesure les discriminations fondées sur l'origine supposée et le genre dans l'accès aux biens et services sur deux plateformes collaboratives françaises de premier plan : Leboncoin et BlaBlaCar.

Le choix s'est porté sur ces deux plateformes car elles offrent la possibilité d'étudier les deux côtés du marché (offre et demande), des interfaces facilitant la récolte de nombreuses données, et proposent plusieurs dispositifs (réservation automatique, pseudonymisation, paiement sécurisé...) susceptibles d'avoir un impact sur les mécanismes de discriminations.

L'étude repose sur deux méthodologies complémentaires :

- un *scraping* ou récolte automatique des données réalisé sur les deux plateformes.
- une expérimentation à partir de profils fictifs (méthode dite du *testing*).

BLABLACAR : LES CONDUCTEURS DÉFAVORISÉS PAR RAPPORT AUX CONDUCTRICES, LES PASSAGERS D'ORIGINE MINORITAIRE DISCRIMINÉS

RÉSULTATS DU SCRAPING

L'analyse des données de BlaBlaCar suggère que les conducteurs ayant un prénom à consonance maghrébine ou africaine (conducteurs d'origine minoritaire) ont un comportement différent des conducteurs majoritaires. Les premiers proposent des prix en moyenne plus élevés, toutefois, lorsque ces différences de prix sont contrôlées, les conducteurs minoritaires reçoivent moins de passagers et un revenu plus faible de 15 % en moyenne par voyage, par rapport aux conducteurs majoritaires¹.

Les conductrices demandent des prix inférieurs aux hommes mais ceci ne semble pas avoir d'effet sur le taux d'occupation de leurs voitures.

TESTING DE CONDUCTEURS FICTIFS

Dans le cadre du *testing*, les conducteurs fictifs d'origine minoritaire reçoivent moins de messages d'intérêt, mais cet effet négatif n'est pas statistiquement significatif².

Les résultats de l'expérimentation ne permettent donc pas d'infirmer ou confirmer les résultats du *scraping* concernant une éventuelle discrimination fondée sur l'origine apparente du conducteur.

En revanche, le *testing* met en évidence une différence significative selon le genre du conducteur. Les conductrices ont en effet une probabilité significativement plus élevée de recevoir des messages (+0,6 messages) et des réservations (+0,7 réservations) que les conducteurs. Cet effet positif envers les femmes n'apparaît pas dans les analyses des données du *scraping*, ce qui souligne

l'importance de compléter les méthodes observationnelles (*scraping*) avec les méthodes expérimentales (*testing*).

LEBONCOIN : UNE FORTE HÉTÉROGÉNÉITÉ POUR LES VENDEURS, UNE DISCRIMINATION IMPORTANTE ENVERS LES ACHETEURS D'ORIGINE MINORITAIRE

RÉSULTATS DU SCRAPING

L'analyse des données du *scraping* suggère que les vendeurs d'origine minoritaire connaissent un temps d'attente plus long pour vendre leurs biens, et donc qu'ils reçoivent moins de contacts de la part des acheteurs que les vendeurs d'origine majoritaire.

TESTING DE VENDEURS FICTIFS

Cependant, les résultats du *testing* révèlent que les vendeurs minoritaires reçoivent en moyenne le même nombre de contacts que les vendeurs majoritaires. Toutefois, ce résultat cache une forte disparité selon la catégorie du bien étudié. Enfin, certaines procédures ou démarches des utilisateurs sur les plateformes dont on aurait pu penser qu'elles participent à la réduction de discrimination éventuelle observée – tels que l'usage de pseudonyme ou encore l'adoption du paiement sécurisé – ne modifient pas le nombre de contacts.

TESTING D'ACHETEURS FICTIFS

Le *testing* montre qu'un acheteur d'origine minoritaire a une probabilité plus faible de recevoir une réponse de la part des vendeurs réels : environ 17 points de pourcentage de moins qu'un acheteur majoritaire. Sur ce côté du marché, la « *pseudonymisation* » proposée par la plateforme est performante pour contrer cet effet discriminatoire : le taux de réponse des acheteurs majoritaires et ceux avec pseudonyme sont statistiquement identiques.

De plus, cet effet discriminatoire à l'encontre des minorités est plutôt homogène quelle que soit la catégorie de bien étudiée, à l'inverse de ce qui est observé du côté des vendeurs.

Cette étude constitue une première étape dans l'étude des discriminations sur les plateformes numériques en France. Les résultats montrent que les discriminations sont présentes dans les échanges sur les plateformes numériques. Toutefois, des différences sont mises en évidence entre les résultats obtenus à partir de l'analyse de l'activité en ligne (*scraping*) et les résultats des expérimentations (*testing*). Ces résultats soulèvent de nombreuses questions qui méritent d'être investiguées, en poursuivant les expérimentations.

INTRODUCTION

L'économie numérique occupe une place croissante dans notre société. Des secteurs entiers de l'économie se retrouvent bouleversés par l'apparition de nouvelles entreprises qui concurrencent les acteurs historiques. Ainsi, 21 % des ventes de livres s'effectuent aujourd'hui sur Internet³. Au moins 40 % des locations immobilières se font dorénavant grâce à des plateformes numériques⁴. Or le fonctionnement de ce nouveau pan de l'économie reste souvent méconnu et ouvre des interrogations quant aux conséquences de son développement. Les plateformes digitales d'intermédiation sont des plateformes de mise en relation au cadre souple qui jouent uniquement un rôle d'intermédiation entre les utilisateurs. Leur intervention est minimale : la personne qui fait l'offre sélectionne elle-même le demandeur parmi les réponses recueillies à son annonce, après avoir fixé le prix des prestations ou du bien.

En modifiant les services proposés et les modalités de contact entre utilisateurs, ces nouvelles plateformes peuvent être génératrices de discriminations, mais pourraient aussi les limiter. D'un côté, les utilisateurs de ces plateformes pourraient davantage mettre en œuvre des biais inconscients et des préjugés que, par exemple, des recruteurs sur le marché du travail. Ces derniers étant plus sensibilisés à la question des discriminations et à leur prévention, ils devraient effectuer leur sélection davantage sur la base de l'évaluation des compétences des candidats et moins sur la base de stéréotypes. De plus, ces plateformes facilitent considérablement le travail de recherche par les utilisateurs, qui ont accès de manière simple à un grand nombre d'annonces de biens et services. Ceci peut donner aux utilisateurs la possibilité de choisir d'interagir avec certaines personnes plutôt que d'autres et donc favoriser la discrimination.

D'un autre côté, une grande partie de ces plateformes mettent en relation des utilisateurs pour acheter un bien ou un service. Cette interaction ne se fait que pendant un laps de temps très court et implique moins d'enjeux que, par exemple, un recrutement sur le marché du travail. Dans le cadre très spécifique des plateformes en ligne, les utilisateurs pourraient donc faire face à une discrimination moindre de la part d'autres utilisateurs. Il est aussi à noter que, grâce à la possibilité de l'anonymat des utilisateurs ou à l'utilisation d'algorithmes⁵, le niveau de discrimination pourrait être réduit sur ces plateformes d'intermédiation.

C'est dans ce contexte que l'étude propose, en réponse à l'appel d'offre du Défenseur des droits, de mesurer et évaluer l'existence potentielle de discriminations sur des plateformes numériques françaises.

L'étude repose sur l'observation de deux plateformes :

- BlaBlaCar est une plateforme de covoiturage permettant de connecter des voyageurs et des conducteurs ayant des places libres dans leurs voitures.
- Leboncoin est une plateforme permettant aux utilisateurs de vendre et d'acheter divers types de biens (ameublement, maroquinerie, jeux vidéo mais aussi des voitures ou même des biens immobiliers).

MÉTHODOLOGIE

Dans la littérature concernant la discrimination sur les plateformes collaboratives, deux méthodes ont principalement été utilisées pour mettre en évidence les discriminations à l'encontre des groupes minoritaires :

- le *scraping*, méthode consistant à récupérer les données publiquement disponibles sur le site internet concerné.
- le *testing* ou l'expérimentation, méthode qui consiste à créer des profils fictifs d'utilisateurs.

Le projet propose d'étudier la discrimination en raison de l'origine supposée et du genre sur les plateformes Leboncoin et BlaBlaCar en combinant ces deux méthodes.

LE SCRAPING

La collecte de données sur les sites concernés permet généralement d'accéder à une quantité de données plus importante que l'expérimentation, au prix d'une relation de causalité plus difficile à établir. En effet, comme les chercheurs ne manipulent pas directement les informations perçues par les utilisateurs, il est difficile d'établir que deux profils de vendeurs ou de conducteurs sont identiques en tous points à l'exception de l'appartenance au groupe discriminé. Malgré cela, et en raison de sa relative simplicité de mise en place, la plupart des études précédentes étudiant la discrimination sur les plateformes collaboratives s'appuient sur le *scraping*.

Un *scraping* systématique des offres postées sur le site de BlaBlaCar a été effectué lors de l'étude. Les données ont été collectées de mars 2020 à juin 2021. Au total, 18 121 observations exploitables ont été recueillies, chacune représentant un trajet de ou vers Paris à partir des 18 plus grandes villes françaises. L'objectif de cette approche est d'obtenir une estimation de l'effet des

caractéristiques des conducteurs sur les prix et le taux de remplissage des voitures. Par ailleurs, ces informations permettent d'identifier les trajets les plus populaires.

Une collecte régulière des annonces sur Leboncoin a également été effectuée. Trois *scrapings* ont été réalisés en janvier, février et avril 2020 et ont permis d'observer plus de 300 000 annonces dans 21 catégories de biens différentes, en récupérant des données sur les annonces des biens (prix, photos, location, texte). Au-delà de l'observation du prix du bien, la durée de la mise en vente du bien, quantifiée en nombre de jours entre la publication de l'annonce et son retrait du site (qui suppose la vente du bien)⁶, a été étudiée.

Pour évaluer la discrimination, il est nécessaire d'associer chaque profil à un groupe d'appartenance (minoritaire ou majoritaire). En se basant sur les prénoms déclarés par les conducteurs et vendeurs, et à partir d'une classification de prénoms largement utilisée dans la littérature⁷, chaque annonce a été associée à une origine supposée du conducteur ou vendeur. Les prénoms ont été classés en trois catégories différentes : prénom d'origine européenne (majoritaire), d'origine maghrébine ou africaine (minoritaire), ou d'origine inconnue. Une opération semblable a été réalisée pour déterminer le genre des conducteurs sur BlaBlaCar.

LE TESTING

Le *testing* par correspondance consiste à créer des profils fictifs d'utilisateurs soit pour envoyer des requêtes d'information aux vrais utilisateurs, soit pour envoyer des demandes d'achat de biens ou de services aux vrais utilisateurs de ces plateformes. Cette expérimentation permet de créer des profils fictifs semblables en tous points, sauf l'appartenance à un groupe pouvant être discriminé.

Par rapport au *scraping*, elle permet donc de s'assurer que l'appartenance à un groupe minoritaire est la seule source possible de différence de traitement des utilisateurs de la plateforme.

Il convient de souligner que ces profils fictifs sont, par définition, vierges de toute réputation ou historique de transactions lorsqu'ils apparaissent sur la plateforme.

Au total 983 annonces fictives ont été postées sur BlaBlaCar, sur les 16 routes les plus fréquentées⁸, entre le 15 mai 2021 et le 15 août 2021. Un algorithme choisissait aléatoirement un conducteur et un trajet parmi une banque de 139 profils de conducteurs fictifs. Ces profils sont similaires entre eux, à l'exception du genre et/ou de l'origine supposée du conducteur, facilement identifiables via le prénom et la photo enregistrés sur le profil et visibles des passagers en recherche d'un trajet. La mesure de la discrimination se fait en comparant les nombres de messages et de réservations reçus par les différents conducteurs fictifs pour un même trajet.

Sur l'autre côté du marché, 1 930 demandes de passagers fictifs ont été envoyées auprès de 965 conducteurs choisis au hasard sur les 11 routes les plus fréquentées, entre le 23 septembre 2021 et le 2 février 2022. Une éventuelle discrimination en raison du genre ou de l'origine supposée était testée. Chacun des conducteurs (réels) recevait ainsi deux demandes de précision concernant le voyage. Ces demandes étaient envoyées à intervalles très rapprochés et étaient en tous points similaires, mis à part le genre des passagers fictifs, ou leur origine (le conducteur recevait une demande d'un homme et d'une femme, ou d'une personne ayant un prénom d'origine européenne et un prénom d'origine maghrébine ou africaine). Les écarts de taux de réponse que les passagers fictifs reçoivent concernant leur demande permettent de mesurer la discrimination.

Sur Leboncoin, les profils des vendeurs fictifs sont classés en trois grandes catégories selon leurs prénoms : Minorités, Majorités et Pseudonymes.

Au total, 332 profils différents ont été créés : 197 profils de vendeurs fictifs d'origine maghrébine et africaine, 104 profils de majoritaires et 31 profils ayant des pseudonymes. Au total, 1 233 annonces ont été postées entre le 19 avril 2021 et le 29 juin 2022 pour la vente de biens accessibles et fréquents (relevant des catégories Ameublement, Consoles et jeux vidéo, Électroménager, Équipement bébé et Vélos).

Le mesure de la discrimination se fait en comparant le nombre total de contacts (par SMS, message vocal ou email) que les vendeurs fictifs reçoivent concernant leurs biens.

Sur l'autre côté du marché, 882 demandes d'informations d'acheteurs fictifs ont été envoyées auprès de 312 vendeurs réels choisis au hasard sur les biens en vente sur la plateforme, entre le 29 juin 2022 et le 12 juillet 2022, selon le même protocole que pour BlaBlaCar. Les demandes concernant la disponibilité du bien en vente étaient un simple prétexte pour stimuler une réponse de l'acheteur. Les différences de taux de réponse que les acheteurs fictifs reçoivent concernant leur demande permettent de mesurer la discrimination.

RÉSULTATS

LES DISCRIMINATIONS EN RAISON DU GENRE ET DE L'ORIGINE SUPPOSÉE SUR LA PLATEFORME BLABLACAR

RÉSULTATS DU SCRAPING

Sur la base des données récupérées par le *scraping*, les hommes, ainsi que les conducteurs ayant un prénom d'origine minoritaire, demandent des prix plus élevés en moyenne pour le même trajet que les conducteurs majoritaires, l'effet étant particulièrement important (environ 4 % plus cher) pour cette deuxième catégorie. Le taux de remplissage des voitures conduites par des hommes, n'est pas différent de celui des femmes. En revanche, les conducteurs d'origine maghrébine et africaine semblent avoir plus de mal à remplir leurs véhicules que les conducteurs majoritaires, même en contrôlant les écarts de prix⁹. Ceci se traduit par un revenu (prix multiplié par le nombre de passagers) plus faible pour ces conducteurs d'environ 1 euro par voyage, malgré un prix plus élevé. Au total, les conducteurs d'origine maghrébine et africaine affichent 15 % de revenu en moins que les conducteurs majoritaires pour des trajets équivalents (graphique 1).

Les femmes sont minoritaires parmi les utilisateurs de BlaBlaCar : selon les données du *scraping* elles représentent environ 30 % des conducteurs et 44 % des passagers. Le *scraping* suggère que les conductrices demandent des prix inférieurs aux hommes mais bénéficient néanmoins d'un taux d'occupation similaire à leurs homologues masculins.

Ces résultats suggèrent la présence de discrimination fondée sur l'origine supposée sur le site de BlaBlaCar, les conducteurs d'origine maghrébine et africaine recevant

des revenus plus faibles pour des offres de trajet semblables (dans la limite des variables explicatives incluses dans la régression). Il ne semble pas y avoir de discrimination fondée sur le genre (graphique 2).

RÉSULTATS DU TESTING : DISCRIMINATION ENVERS LES CONDUCTEURS

Cette démarche teste de manière expérimentale l'effet du genre et de l'origine supposée de conducteurs fictifs sur le remplissage de leurs voitures. Les variables d'intérêt ici sont les demandes de réservations et autres retours (message, demande de précision) que peuvent recevoir ces conducteurs.

Les femmes bénéficient d'un traitement avantageux qui se traduit non seulement dans le nombre de demandes de réservations fermes, mais aussi dans des notifications d'intérêt telles que les demandes de précision sur le trajet qu'elles reçoivent. Chaque annonce a reçu en moyenne 1,4 demande de réservation et 2,5 signes d'intérêt (demandes de précision ou réservations effectives)¹⁰.

En prenant en compte les caractéristiques des annonces et des routes, il est constaté que les profils avec des photos et des prénoms masculins sont discriminés négativement. En effet, les femmes reçoivent en moyenne plus de demandes de précisions (0,6) avant réservation, de réservations (0,4) et de signes d'intérêt (0,9) que les hommes (graphique 3).

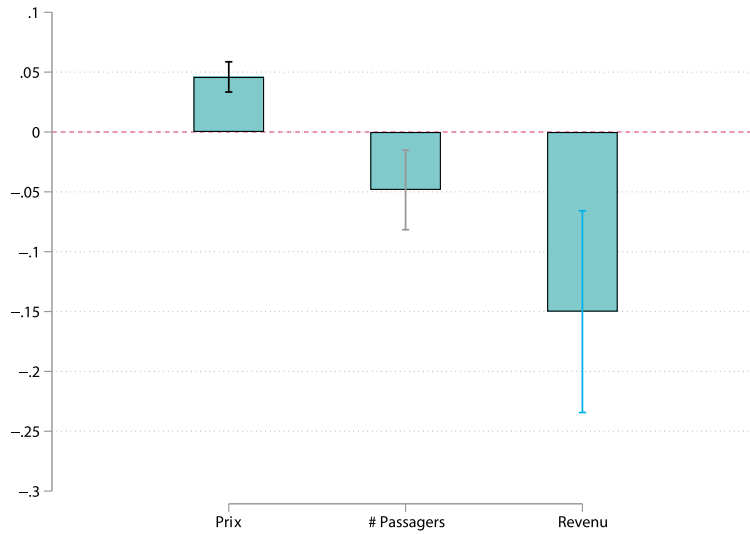
L'origine supposée du conducteur fictif n'a en revanche pas d'effet significatif direct, à l'inverse de ce qui a été observé dans les résultats du *scraping*. En effet, les données récoltées par *scraping* sur BlaBlaCar suggèrent que les conducteurs d'origine minoritaire ont 0,048 passagers de moins dans leurs voitures.

En prenant la mesure analogue dans l'expérimentation (le nombre de réservations par annonce fictive) les conducteurs de ce groupe reçoivent 0,045 demande de plus ce qui contredit le résultat du *scraping*.

Ce coefficient est cependant estimé avec beaucoup d'imprécision et les autres variables d'intérêt ont un coefficient négatif (graphique 4).

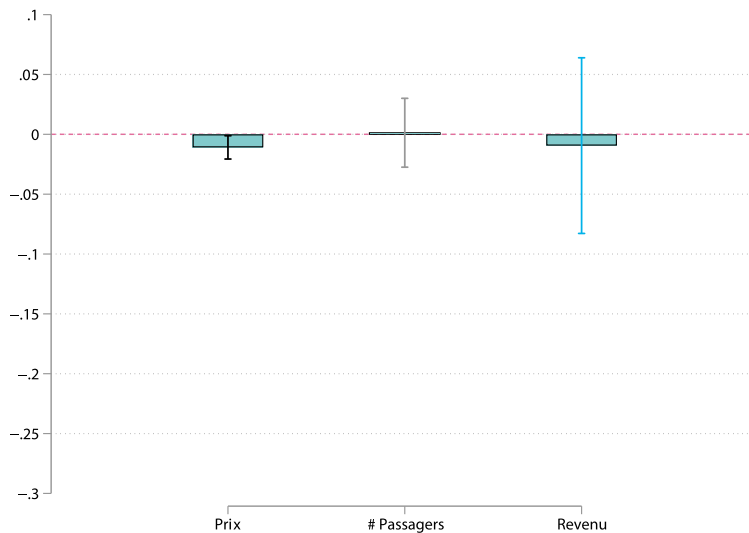
GRAPHIQUE 1

Effet d'avoir un prénom à consonance africaine ou maghrébine sur le prix moyen, le nombre de passagers ou le revenu moyen du trajet – données du *scraping* sur BlaBlaCar



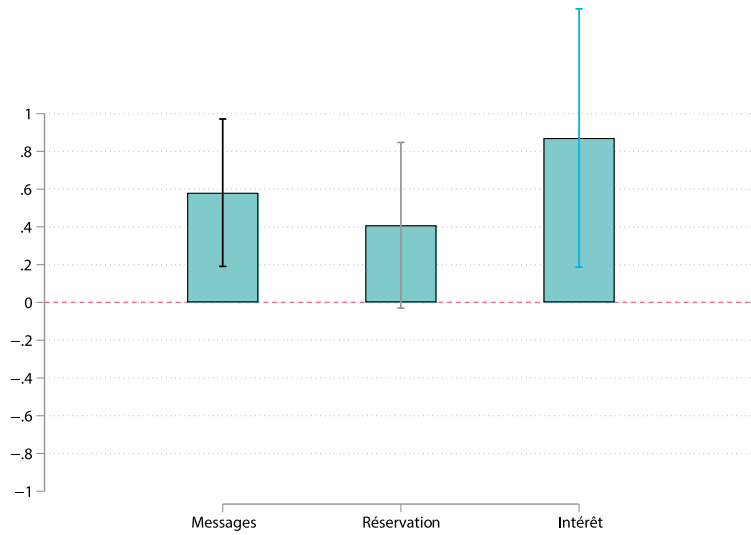
GRAPHIQUE 2

Effet d'avoir un prénom féminin sur le prix moyen, le nombre de passagers ou le revenu moyen du trajet – données du *scraping* sur BlaBlaCar



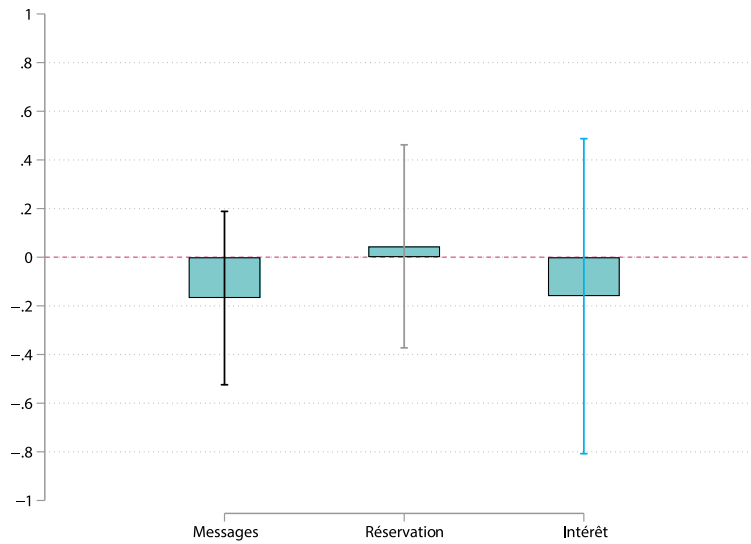
GRAPHIQUE 3

Effet d'avoir un prénom féminin sur le nombre de messages, de réservations et signes d'intérêt reçus – données du *testing* sur BlaBlaCar



GRAPHIQUE 4

Effet d'avoir un prénom à consonance africaine ou maghrébine sur le nombre de messages, de réservations et signes d'intérêt reçus – données du *testing* sur BlaBlaCar



RÉSULTATS DU TESTING : DISCRIMINATION ENVERS LES PASSAGERS

Des demandes de renseignement ont ensuite été envoyées aux conducteurs en provenance de faux profils de passagers différant sur le genre et l'origine supposée uniquement : le taux moyen de réponse à ces demandes est de 63 %.

Les résultats suggèrent que les passagers d'origine minoritaire ont une probabilité plus faible de recevoir une réponse à leur demande, d'environ 4 points de pourcentage.

Par ailleurs, l'étude de l'interaction entre genre et origine supposée du passager ne montre un effet négatif que pour les hommes - les femmes d'origine minoritaire ont la même probabilité de recevoir une réponse que les passagères d'origine majoritaire.

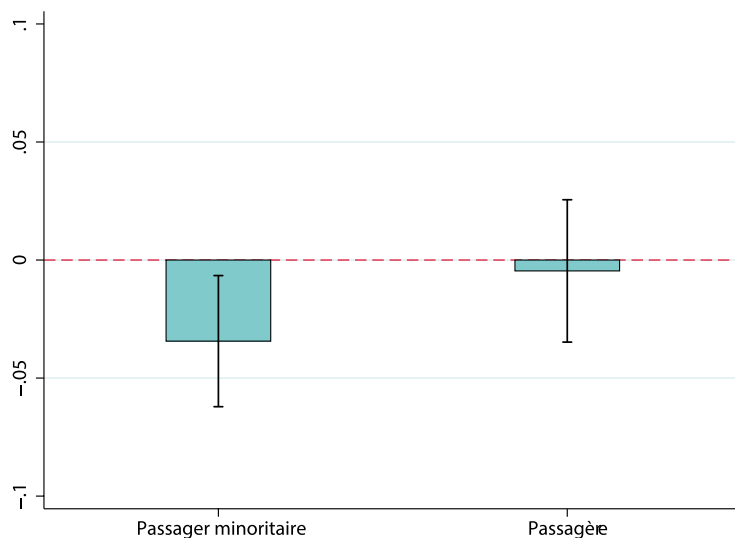
Les résultats du *testing* concernant l'interaction entre passagers fictifs et conducteurs réels suggèrent qu'il n'y a pas

de différence de taux de réponse entre passagers fictifs d'origine minoritaire et majoritaire lorsque le conducteur est lui-même d'origine minoritaire. Le genre du conducteur n'induit pas de différences importantes ou significatives. Les conductrices semblent donc agir de la même façon que les conducteurs.

En outre, le *testing* permet d'analyser la provenance des demandes : nous observons que les propositions de trajet émanant de conductrices suscitent davantage d'intérêt aussi bien auprès des passagers femmes que des passagers hommes. Les conductrices d'origine européenne semblent, plus particulièrement, attirer des demandes émanant de passagers femmes. De l'autre côté du marché, les résultats ne révèlent pas de discrimination en raison du genre à l'encontre des passagers. Nous ne trouvons pas non plus de préférence de la part des conductrices pour les passagères.

GRAPHIQUE 5

Effet d'avoir un prénom à consonance minoritaire ou féminine sur le nombre de réponses à une demande de réservation – données du *testing* sur BlaBlaCar



LES DISCRIMINATIONS EN RAISON DU GENRE ET DE L'ORIGINE SUPPOSÉE SUR LA PLATEFORME LEBONCOIN

RÉSULTATS DU SCRAPING

Un sous-échantillon de 10 catégories a été retenu. Pour chacune d'elles, les vendeurs ont été répartis en deux groupes (minoritaire, majoritaire) en fonction de leur nom (excluant les annonces vendues sous pseudonymes ou par des vendeurs dont le prénom n'a pas pu être apparié). Les catégories qui n'exigent pas des utilisateurs de référencer les caractéristiques de leurs biens (modèle, marque...) ont été exclues pour assurer la comparabilité des offres⁴¹.

L'analyse consiste à étudier les prix affichés pour des biens similaires ainsi que la durée de vie sur le site du bien entre les deux groupes de vendeurs (mineurs et majoritaires).

Pour les catégories Motos, Montres et Bijoux, et Vélos, les prix des annonces postées par les utilisateurs du groupe majoritaire sont plus élevés que pour les mineurs, à biens équivalents (c'est-à-dire en ajustant pour les caractéristiques observables du bien et de l'annonce). Pour les catégories Électroménager, Équipement bébé, Consoles et jeux vidéo et Chaussures, la relation est inversée : les prix des annonces postées par les utilisateurs majoritaires sont plus faibles que pour les mineurs. Les résultats indiquent aussi que, pour 8 catégories parmi les 10 étudiées, les biens du groupe majoritaire restent moins longtemps sur le site en ajustant sur les caractéristiques observables du bien et de l'annonce.

Ces résultats contrastés peuvent signifier l'existence de la discrimination sur cette plateforme ou, simplement, que les biens postés sont différents sur d'autres dimensions que nous n'observons pas avec les données collectées par le *scraping*. Le *testing* avec les vendeurs fictifs nous permet de contrôler cette possibilité en rendant les biens identiques sur toutes les dimensions.

RÉSULTATS DU TESTING : DES DISCRIMINATIONS ENVERS LES VENDEURS

L'étude ne révèle pas de discrimination envers les vendeurs mineurs. Les vendeurs fictifs mineurs⁴² reçoivent en moyenne 0,7 contact de plus que les vendeurs majoritaires mais cet effet estimé n'est toutefois pas statistiquement significatif avec notre échantillon. Les vendeurs avec pseudonyme reçoivent moins de contacts (- 0,52) pour le bien en comparaison avec les vendeurs majoritaires mais cet effet n'est pas non plus statistiquement significatif (graphique 6). Il n'est pas observé non plus d'impact du groupe mineur sur la probabilité d'avoir reçu au moins une requête en comparaison avec les vendeurs majoritaires.

Ces analyses montrent que, en moyenne, les vendeurs mineurs et les vendeurs avec pseudonyme reçoivent le même nombre de requêtes pour leurs biens que les vendeurs majoritaires (graphique 6).

L'étude s'attache ensuite à analyser les variations selon la catégorie du bien proposé par le vendeur.

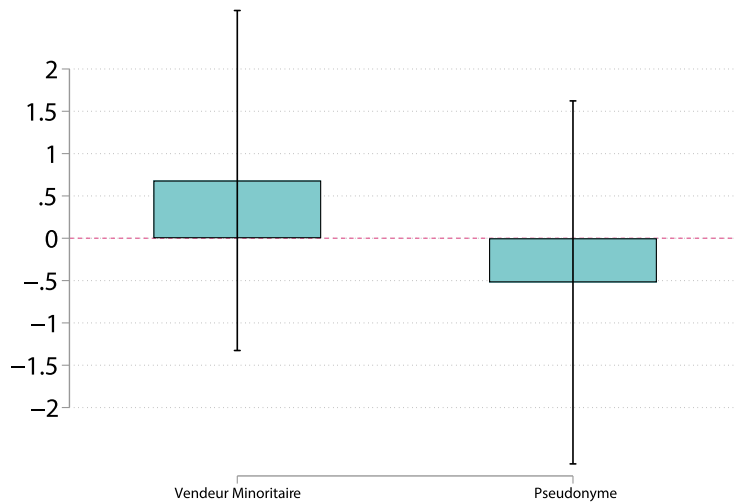
L'impact moyen de l'appartenance à un groupe⁴³ sur l'ensemble de l'échantillon masque une forte hétérogénéité selon la catégorie du bien (graphique 7). Pour les catégories Ameublement et Équipement bébé, un effet positif et statistiquement significatif lié aux vendeurs mineurs a été retrouvé. En d'autres termes, les vendeurs perçus comme mineurs reçoivent davantage de contacts pour ces catégories. Cet effet s'inverse pour les trois autres catégories - Électroménager, Console et Jeux Vidéo et Vélos - pour lesquelles les minorités reçoivent moins de demandes. La réduction de la taille de l'échantillon réduit la précision des estimations. Cependant, si les effets ne sont pas toujours statistiquement significatifs lorsque l'on considère chaque catégorie séparément, nous pouvons en conclure que l'impact est statistiquement différent à travers les catégories. Ainsi, la non-existence de discrimination en fonction de l'origine supposée du vendeur sur l'ensemble des biens semble masquer des effets différenciés selon les marchés considérés.

En résumé, les comportements des acheteurs potentiels semblent très différents selon les catégories de biens considérés. Dans le cas des catégories Équipements bébé et Ameublement, les vendeurs minoritaires reçoivent davantage de demandes de contacts alors qu'ils en reçoivent moins pour les Vélos, l'Électroménager et les Consoles/jeux vidéo.

Ces résultats conduisent à conclure à l'absence de discrimination systématique à l'encontre des minorités proposant leurs biens sur la plateforme. Toutefois, nos résultats n'excluent pas l'existence de comportements différenciés des acheteurs potentiels en fonction de l'origine supposée des vendeurs sur certains marchés.

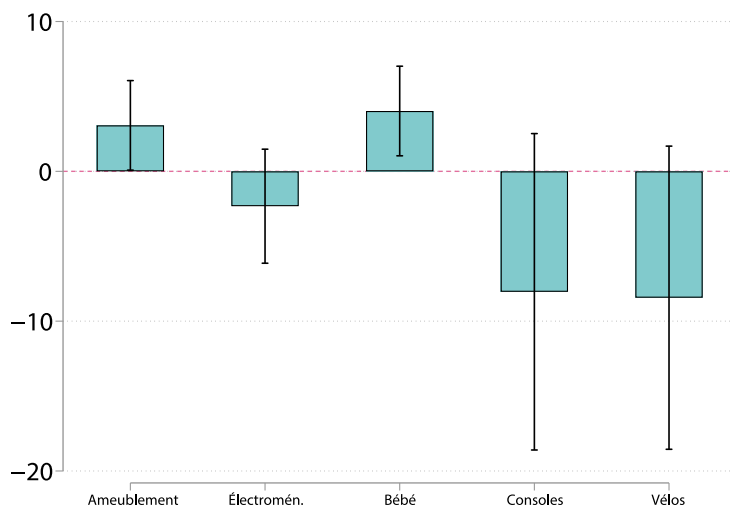
GRAPHIQUE 6

Effet d'avoir un prénom à consonance minoritaire ou un pseudonyme sur le nombre de contacts de la part d'acheteurs – données du *testing* sur Leboncoin



GRAPHIQUE 7

Effet d'avoir un prénom à consonance minoritaire sur le nombre de contacts de la part d'acheteurs en fonction de la catégorie de biens – données du *testing* sur Leboncoin



RÉSULTATS DU TESTING : DES DISCRIMINATIONS ENVERS LES ACHETEURS

En revanche, le *testing* met en évidence la présence d'une discrimination à l'encontre des acheteurs minoritaires. La différence de réponses aux demandes d'informations d'acheteurs minoritaires par rapport à celles émanant d'acheteurs majoritaires est très importante : en moyenne, 82 % des sollicitations envoyées par des acheteurs fictifs ont reçu une réponse du vendeur mais le taux de réponse pour les acheteurs minoritaires est plus faible de 17 points de pourcentage par rapport aux acheteurs majoritaires (graphique 8). En revanche, le taux de réponse entre le groupe d'acheteurs avec pseudonyme et le groupe majoritaire n'est pas statistiquement différent.

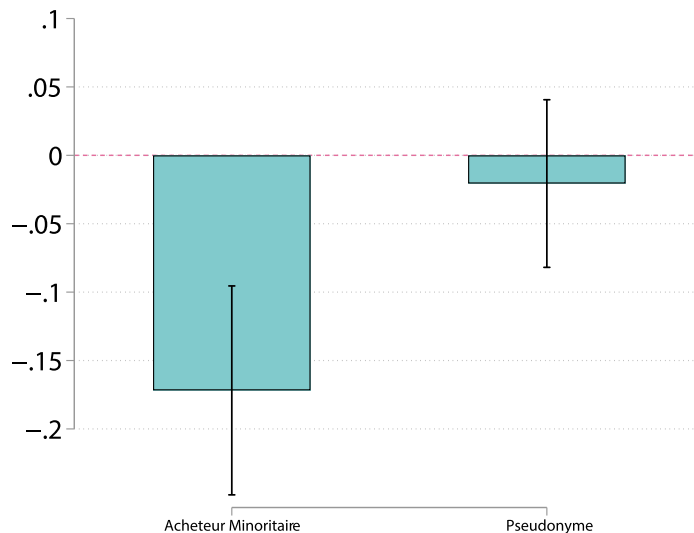
Ce résultat indique que si un acheteur minoritaire adoptait un pseudonyme masquant son origine supposée, il pourrait augmenter le taux de réponse de la part des vendeurs de manière significative.

Les analyses¹⁴ mettent en évidence des effets négatifs à l'encontre des potentiels acheteurs minoritaires pour toutes les catégories de biens (graphique 9). Toutefois, ces différences sont statistiquement significatives pour deux catégories seulement (Ameublement et Vélos). Ces résultats ne permettent pas de conclure que l'impact de la discrimination est différent selon la catégorie du bien, à l'inverse de ce que l'on trouve côté offre.

Il s'avère donc que la discrimination à l'encontre des acheteurs minoritaires est un phénomène qui se manifeste de manière plus ou moins homogène sur cette plateforme.

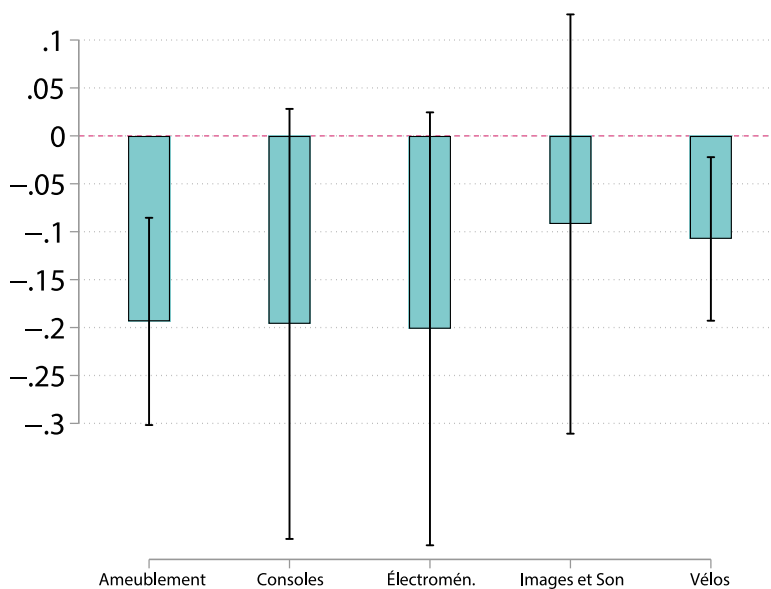
GRAPHIQUE 8

Effet d'avoir un prénom à consonance minoritaire sur le nombre de réponses de la part des vendeurs – données du *testing* sur Leboncoin



GRAPHIQUE 9

Effet d'avoir un prénom à consonance minoritaire sur le nombre de réponses de la part des vendeurs, en fonction de la catégorie de biens – données du *testing* sur Leboncoin



CONCLUSION

Cette étude constitue une étape importante dans l'étude des discriminations sur les plateformes numériques en France. Les résultats montrent que les plateformes numériques ne sont pas protégées contre les discriminations.

En premier lieu, des différences importantes sont mises en évidence entre les résultats obtenus à partir de l'analyse de l'activité en ligne (*scraping*) et les résultats de l'expérimentation (*testing*). Sur la plateforme BlaBlaCar, les résultats du *scraping* et de l'expérimentation ne sont pas concordants : les résultats issus du *scraping* indiquent une absence de discrimination liée au genre alors que le *testing* montre la présence d'un traitement préférentiel en faveur des conductrices. De même, le *scraping* semble montrer une discrimination envers les conducteurs minoritaires, alors que celle-ci n'est pas détectée avec le *testing*.

Il en est de même pour la plateforme Leboncoin. Si les résultats du *scraping* montrent que pour la grande majorité des catégories, les biens des vendeurs du groupe majoritaire sont plus rapidement vendus que ceux du groupe minoritaire, l'expérimentation ne corrobore pas complètement ces résultats. Le *testing* montre en particulier que les comportements des acheteurs potentiels semblent très différents selon les catégories de biens considérés.

Ce résultat suggère de mieux connaître les différences de comportements selon les catégories de biens avant d'élaborer des politiques pour réduire les discriminations en ligne. De plus, si notre étude suggère que la mise en place d'une « *pseudonymisation* » complète sur Leboncoin du côté des vendeurs pourrait limiter les comportements discriminatoires, cette mesure pourrait aussi remettre en cause l'efficacité du marché dans certains cas.

En revanche, les résultats révèlent de la discrimination à raison de l'origine à l'encontre des passagers sur la plateforme BlaBlaCar et des acheteurs sur Leboncoin. Sur BlaBlaCar, il est observé une discrimination importante de

la part des conducteurs envers les passagers masculins ayant des prénoms d'origine maghrébine ou africaine. Sur Leboncoin, les acheteurs minoritaires ont beaucoup moins de chance de recevoir une réponse des vendeurs à leur demande d'information. Le recours à la « *pseudonymisation* » peut s'avérer efficace pour réduire la discrimination. Il est aussi important de rappeler que l'expérimentation sur Leboncoin porte principalement sur les premiers échanges et les demandes de renseignement et non l'achat effectif du bien. Ainsi, il se peut qu'une discrimination intervienne lors des étapes ultérieures du processus et non lors du premier contact. Si cette question reste pertinente pour Leboncoin, il convient de noter que dans le cas de BlaBlaCar, les estimations restent similaires, que l'on observe les demandes d'informations ou bien directement les réservations.

Comme évoqué précédemment, des études ont montré l'existence de discrimination via les données de *scraping* et via les expérimentations, même si pour cette deuxième méthode, il n'y a pas de consensus clair sur l'amplitude de la discrimination sur les marchés en ligne, car celle-ci dépend fortement du contexte. Si les résultats du *scraping* mené sur Leboncoin et BlaBlaCar semblent convergents avec les études précédentes, ce n'est pas le cas des résultats issus des expérimentations par *testing*.

Ces résultats suggèrent que les mécanismes en jeu sont complexes, et que les travaux pour documenter ces questions doivent être poursuivis à partir de méthodologies diverses et complémentaires (*testing*, *scraping*...). L'impact de la pseudonymisation et l'impact d'une appartenance minoritaire varient selon le côté du marché (vendeur ou acheteur) et selon la catégorie de biens considérés.

Ces résultats soulèvent de nombreuses questions qui méritent d'être investiguées, en poursuivant les analyses, afin de continuer à mettre en lumière et prévenir les discriminations sur ces plateformes qui occupent une part toujours plus importante dans nos sociétés.

BIBLIOGRAPHIE

Ajay A., Lacetera N., Lyons E. (2016). "Does standardized information in online markets disproportionately benefit job applicants from less developed countries?", *Journal of international Economics*, 103: 1-12.

Ian A., Banaji M., Jolls C. (2015). "Race effects on eBay", *The RAND Journal of Economics*, 46 (4): 891-917.

Becker GS. (1957). The economics of discrimination, *University of Chicago press*.

Brutel C. (2016). « La localisation géographique des immigrés : une forte concentration dans l'aire urbaine de Paris ». *Rapport technique, Insee Premiere* 1591.

Card D., DellaVigna S., Funk P., Iriberry N. (2020). "Are Referees and Editors in Economics Gender Neutral?", *The Quarterly Journal of Economics* 135 (1): 269-327.

Chapelle G., Eyméoud JB. (2022). "Can big data increase our knowledge of local rental markets ? A dataset on the rental sector in France", *PLOS ONE* 17 (1): 1-21.

Cody C., Diamond R., Hall JV., List JA., Oyer P. (2021). "The gender earnings gap in the gig economy: Evidence from over a million rideshare drivers", *The Review of Economic Studies* 88 (5): 2210-2238.

Ruomeng C., Li J., Zhang DZ. (2020). "Reducing discrimination with reviews in the sharing economy: Evidence from field experiments on Airbnb", *Management Science* 66 (3): 1071-1094.

Doleac J., Stein LCD. (2013). "The visible hand: Race and online market outcomes", *The Economic Journal* 123 (572): F469-F492.

Edelman B., Luca M. (2014). "Digital discrimination: The case of Airbnb.com", *Harvard Business School NOM Unit Working Paper*, n° 14-054.

Edelman B., Luca M., Svirsky M. (2017). "Racial discrimination in the sharing economy: Evidence from a field experiment", *American Economic Journal: Applied Economics* 9 (2): 1-22.

Mehdi F., Hammond RG., Pénard T. (2019). "What drives pricing behavior in peer-to-peer markets? Evidence from the carsharing platform BlaBlaCar", *Information Economics and Policy* 48: 15-31.

NOTES

¹ ie deux conducteurs affichant les mêmes prix pour un même trajet, le conducteur d'origine maghrébine ou africaine a un taux de remplissage et donc un revenu plus faible que le conducteur d'origine européenne.

² La significativité statistique désigne le seuil à partir duquel les résultats d'un test sont jugés fiables, ou autrement dit, l'échantillon utilisé est suffisamment représentatif de la population globale et de son comportement.

³ Chiffres-clés du secteur du livre 2018-2019, Publication du ministère de la Culture.

⁴ Chapelle G., Eyméoud JB. (2022) *Can big data increase our knowledge of local rental markets? A dataset on the rental sector in France*. PLoS ONE 17(1).

⁵ L'utilisation d'algorithmes permettrait par exemple de modifier l'ordre ou la visibilité de certaines annonces ou informations, ou encore de repérer d'éventuels comportements discriminatoires de certains utilisateurs qui seraient invités à modifier leur comportement.

⁶ Cette approche fait l'hypothèse que le vendeur retire son annonce du site une fois le bien vendu.

⁷ Jouniaux L. (2001) « *Les 20 000 plus beaux prénoms du monde* », Hachette éd.

⁸ Les trajets relient Paris à l'une des villes suivantes : Lyon, Lille, Toulouse, Strasbourg, Rennes, Nice, Grenoble, Nantes, Bordeaux, Reims, Clermont-Ferrand, Montpellier, Dijon, Le Havre, Tours, Marseille.

⁹ ie pour deux conducteurs affichant les mêmes prix pour un même trajet, le conducteur d'origine maghrébine ou africaine a un taux de remplissage et donc un revenu plus faible que le conducteur d'origine européenne.

¹⁰ Ainsi les personnes envoyant plusieurs messages ou un message et une demande de réservation ne sont comptées qu'une seule fois.

¹¹ Pour certaines catégories, une annonce d'un bien est seulement représentée par une photo, un titre et un court descriptif, ce qui ne permet pas de s'assurer de la comparabilité des biens et de la durée de l'annonce.

¹² Vendeurs avec un prénom d'origine supposée maghrébine ou africaine subsaharienne.

¹³ Les vendeurs avec pseudonyme ont été regroupés avec les vendeurs majoritaires pour augmenter la puissance statistique et donc améliorer la précision des estimations.

¹⁴ Pour cette analyse, les vendeurs majoritaires et avec pseudonyme ont été regroupés, en comparaison avec les vendeurs minoritaires pour augmenter la puissance statistique Il est cohérent d'effectuer cela car nous ne détectons aucune différence significative entre le taux de réponse des acheteurs majoritaires et ceux avec pseudonyme.

Dans le cadre de ses missions confiées par la Loi organique du 29 mars 2011, le Défenseur des droits soutient des travaux d'études et de recherches intéressant ses différents champs de compétence :

- la défense et la promotion des droits des usagers des services publics ;
- la défense et la promotion des droits de l'enfant ;
- la lutte contre les discriminations et la promotion de l'égalité ;
- le respect de la déontologie par les personnes exerçant des activités de sécurité ;
- l'orientation et la protection des lanceurs d'alerte.

La collection Éclairages se propose de mettre à disposition des spécialistes, des décideurs, des professionnels comme du public le plus large, les synthèses des travaux menés par des équipes de recherche pluridisciplinaires et indépendantes pour le compte de l'institution. Elle a vocation à éclairer le débat public et documenter les enjeux de l'intervention du Défenseur des droits.



Défenseur des droits
RÉPUBLIQUE FRANÇAISE