
Profit vs morality: how unfair is labor market discrimination? Results from a survey experiment

Document de Travail
Working Paper
2019-25

Matthieu Bunel
Élisabeth Tovar



UMR 7235

Economix - UMR7235
Université Paris Nanterre
Bâtiment G - Maurice Allais, 200, Avenue de la République
92001 Nanterre cedex

Email : secretariat@economix.fr



PROFIT VS MORALITY

RESULTS FROM A SURVEY EXPERIMENT

ON DISCRIMINATION¹

Mathieu Bunel

University of New Caledonia and LARJE (EA 3329)
Campus de Nouville BP R4 98851 Nouméa Cedex, France
Mathieu.Bunel@unc.nc

and

Elisabeth Tovar²

Paris Nanterre University and EconomiX (UMR 7235)
200 avenue de la République 92001 Cedex Nanterre, France
elisabeth.tovar@parisnanterre.fr

Highlights

- vignette-based survey experiments reveal causal effects on normative preferences
- normative opinions on discrimination are affected by profit considerations
- customer taste discrimination is more acceptable than screening or statistical discrimination
- moral suasion effects have an impact on normative choices between profit and morality

Abstract

Using an original survey-experimental protocol, we study the normative acceptability of the trade-off between immoral profit (discrimination) and costly morality (non-discrimination). We test the causal influence of three factors: i) the origin of discrimination, ii) the steepness of the morality/profit trade-off and iii) anti-discriminatory moral injunctions. Contrasting with past experimental and attitudinal studies, we find that a significant minority of respondents believe that labor market discrimination is acceptable when morality results in profit loss. We also find that the three tested factors have significant effects on normative opinions. Respondents are more likely to choose profit over morality when discrimination is taste-based than when it is caused by imperfect information. Discrimination's acceptability rises with the cost of non discrimination. Anti-discriminatory moral injunctions sharply reduces the acceptability of profitable discrimination.

Keywords: Discrimination, moral suasion, profit/morality trade-off, vignette survey experiment

JEL classification: J23, J7, J78, C9

Funding: this work was supported by the Agence Nationale de la Recherche Jeune Chercheur 2015 Grant programme

¹ Warm thanks to Sophie Harnay and Régis Lanneau for their help during the administration of survey. Thanks to the participants of the Nouméa workshop on discrimination, Frédéric Chantreuil, Isabelle Lebon and Dominique Meurs. Thank you also to the two EconomiX referees.

² Corresponding author.

1. Introduction

In Becker's first framework (Becker, 1957), discrimination happens on a market if an agent is willing to lose profit to exclude a target group he or she dislikes. This behavior being sub-optimal, in the long run agents with discriminatory preferences should be driven out of the market. Moreover, survey data shows that discrimination is widely condemned in most countries. For example, in France, where discrimination is illegal³, 90% agree that discriminating in the labor market⁴ when an applicant has all the required qualifications is a major ("grave" in French) infraction (CNCDH 2017). Similar results have been found in the literature. In Barr, Lane and Nosenzo (2018), up to 85% respondents declare that an unequal allocation between groups is "inappropriate". In Dickson et al (2018), participants to an experimental game are willing to sacrifice part of their income to sanction the behavior of employers who display in-group favoritism.

However, survey data also shows that a significant proportion of the population declares having been discriminated against. In the French example, 10% of the population reports having been at least once the victim of discrimination when applying to a job (Generations survey, CEREQ, 2017). This ratio rises respectively to 41% and 36% for workers with a North African or Sub Saharan African origin. 16.8% of the European Social Survey (ESS) respondents interrogated in France in the 2008 wave of the ESS reported having been treated with prejudice in the past year because of their ethnic background (26.7% because of their gender; 34.7% because of their age)⁵. Experimental data compounds this attitudinal evidence. Lane's meta-analysis (2016) shows that on average 1/3 of the population display a preferential treatment of members of their own group. Such in-group favoritism is more frequent when groups are formed on an artificial (randomly draw) or socio-geographic basis than on other characteristics (nationality, gender, religion, ethnicity)⁶. Evidence accumulated through testing campaigns in the labor market (Neumark, 2018), the housing market (Oh and Yinger 2015) and the credit market (Ross and Yinger 2002) also shows that discriminatory behaviors persist in most countries and in most markets (Bertrand and Duflo, 2017).

The economic literature proposes two alternative mechanisms to account for rational, long-lasting discriminatory behaviors. First, Welch (1967) and Becker (1971) focus on the discriminatory preferences of the customers or employees of the decision-maker. In this case, the decision-maker faces a trade-off between profit and morality: whatever her or his own preferences, discrimination is now profitable if he or she needs to cater to his or her clients or employees.

Second, discrimination may stem from imperfect information. Arrow (1973) and Phelps (1972) provide a model of statistical discrimination where group membership is used as a signal for unobservable individual characteristics. If a target group suffers from a (real or perceived) quality deficit by comparison with other groups, it may be rational for decision-makers to discriminate against members of this group. In this case, the trade-off is between the cost of acquiring accurate information on the group members' individual characteristics and the risk of dismissing a productive member of the target group in favor of another, less productive, member of a non-targeted group. Depending on the outcome of the trade-off, discrimination can be either profitable or not.

In this paper, our aim is to shed evidence on the seemingly paradoxical fact that discrimination is a widespread phenomenon although most people strongly disagree with it from a moral point of view. We use experimental survey evidence to explore causal effects on both sides of the profit/morality trade-off faced by decision-makers in the labor market.

³ In most countries legislation prohibits discriminatory behaviors, since they violate the equality principles enforced by most constitutions. In France, the article 225 of the Penal Code states that "*is a discrimination any distinction between natural or moral persons on the basis of their origin, sex, family status, pregnancy, physical appearance, particular vulnerability resulting from their economic situation, apparent or known to the author, their surname, place of residence, state of health or loss of autonomy shall constitute discrimination, their disability, genetic characteristics, mores, sexual orientation, gender identity, age, political opinions, trade union activities, ability to express themselves in a language other than French, their true or supposed membership or non-membership, true or supposed, of a particular ethnic group, nation, alleged race or religion.*"

⁴ In this paper, we narrow the focus on labor market discrimination, although similar mechanisms are at play on other markets, such as the housing of the credit market.

⁵ See Valfort (2018) for statistics on European countries.

⁶ Recently, Barr, Lane and Nosenzo (2018) found that 1 out of 6 participants in a lab experiment discriminated against individuals who do not belong to their group when the groups are nationality-based. This rate is 3 out of 10 when the groups are artificial (randomly drawn).

This issue is widely discussed in moral philosophy (see Arneson, 2006, for an in-depth discussion). Although many papers document discriminatory behaviors (see Lang and Lehmann, 2012, Neumark, 2018 for comprehensive surveys), the normative acceptability of discrimination has received little attention so far in the economic literature (Chassonnery-Zaigouche, 2012, Barr, Lane and Nosenzo, 2018). Our goal here is to provide evidence on the factors that affect the *normative acceptability* of the profit/morality trade-off in the eyes of the population.

To do so, we use an original questionnaire-experimental vignette survey depicting a trade-off between supporting a moral (i.e., not discriminatory) but unprofitable behavior and supporting a profitable but discriminatory behavior. We use factorial design to manipulate three contextual factors: the cost of non-discriminatory behaviors, the nature of discrimination and the presence of an anti-discriminatory moral injunction. The experiment was presented between August 2018 and January 2019 to about 1,100 students enrolled in three French universities (Paris Nanterre, Paris Descartes and New Caledonia). By comparing the answers of the groups of respondents randomly assigned to each version of the survey, we provide causal evidence on how the contextual factors affect the normative preference of the respondents in the profit/morality trade-off.

We find three main results. First, all versions of the vignette considered, 40% of the respondents agree with a significant labor market discrimination of the target group. Such a high level of discrimination acceptability is in sharp contrast with previous studies based on direct questioning about the fairness of discriminatory behaviors (Barr, Lane and Nosenzo 2018). It however echoes with subjective data collected in the attitudinal surveys mentioned above. Moreover, we find that the cost of non-discrimination has a significant effect on the normative profit/morality trade-off. Discrimination becomes more morally acceptable when it avoids an important profit loss to the perpetrator. This result echoes those of Zussman (2013) and Tyran and Hedegaard (2018) by pointing that normative preferences, as well as behaviors, take into account the cost of morality. Second, public opinion is also affected by the motive of discrimination: respondents are less prone to accept discrimination when it is based on imperfect information than when it is caused by consumer preferences (by a gap of almost 20 percentage points). Third, the presence of a moral suasion effect, whether egalitarian or compensatory, significantly affects the respondent's moral compass by reducing the acceptability of discrimination.

The article is organized as follows. Section 2 briefly presents the literature on labor market discrimination and moral suasion effects. Section 3 focuses on the vignette methodology and our empirical strategy. Section 4 presents and discusses our results. Section 5 concludes.

2. Literature review

Discrimination is prevalent in many markets, as shown by a vast empirical and experimental literature (for a survey, see Lang and Lehmann, 2012, Bertrand and Duflo 2017, Neumark, 2018). Here, we will focus on the literature on (i) the motives of rational discrimination, (ii) the role of the profit/morality trade-off on discriminatory behaviors and (iii) the influence of moral suasion effects on behaviors and preferences.

2.1. Motives of rational discrimination

Rational discrimination, where the decision-maker has no discriminatory preferences himself or herself, is driven by profit/morality trade-offs: the non-discriminatory behavior is costly for the employer, either because he or she risks losing clients or employees, or because he or she risks hiring low-productivity workers. The literature classically explains rational discrimination through two main mechanisms (Guryan and Charles, 2013).

A first major motive of discrimination is based on the differentiated preferences of employers, customers or employees. In Becker's first framework (1957), discrimination is caused by employers who make a difference between applicants belonging to different groups. As Feld et al. (2016) point out, this differentiated treatment may result either from endophilia (in-group favoritism) towards the members of the employer's own group or from hostility (exophobia) towards the members of a group to which the employer does not belong. Whether stemming from endophilia or exophobia, taste-based discrimination

does not have a rational basis: no explanation is provided for the dis-utility perceived by the employer from contact with members of the discriminated group. This initial taste-based model of discrimination was generalized by Welch (1967) and Becker (1971) to account for the tastes of customers and employees. In this case, even if the employer does not have discriminatory preferences herself/himself, it can be rational for him/her to cater to the preferences of his or her customers and employees. He or she then faces a morality/profit trade-off: if he or she does not follow to the preferences of his or her customers (or employees) he or she risks losing customers (or having difficulties hiring or keeping productive employees) and therefore lose money.

Imperfect information is the second major motive of rational discrimination (Phelps, 1972; Arrow, 1973). When individual productivity characteristics are not observable, employers tend to assign to all members of any given group the actual (or assumed) average productivity of the whole group. Members of targeted groups are discriminated against because hiring decisions are independent from their actual individual characteristics. In this case, there is a profit trade-off between the cost of acquiring accurate information on the group members' individual characteristics and the risk of dismissing a productive member of the target group in favor of another, less productive, member of a non-targeted group. Depending on the trade-off, discrimination can be either profitable or not. This statistical discrimination model may be fueled by prejudiced beliefs about members of a particular group (Phelps, 1972; Arrow, 1973) or by the quality of the productivity signal as perceived by employers who have poor intimate knowledge of the discriminated group (Aigner and Cain, 1977). In the latter case, known as discrimination screening (or rational homophily), there is an asymmetry of information between the quality of the employer's information about members of his or her own ethnic/racial/sex group and the quality of the information about applicants who do not belong to that group (Cornell et al., 1996, Pinkston, 2003).

Our experimental framework is designed to test whether the moral acceptability of discrimination is sensitive to the motive of discriminatory behaviors: client-based, statistical or screening discrimination. A first working hypothesis is that respondents will find discrimination less acceptable when the decision-maker's responsibility is involved. In this case the acceptability of discrimination will be highest when it is justified by clients' tastes and lowest when it results from the decision-maker's sensitivity to rumors or lack of knowledge of discriminated groups [working hypothesis 1] . A second working hypothesis is that respondents take into account the difference in the probability of the loss of profit caused by the different grounds of discrimination [working hypothesis 2]. If this hypothesis is verified, they should find client-based discrimination more acceptable than that resulting from imperfect information (screening and statistical discrimination). Finally, a final hypothesis echoing the work of Feld et al. (2016) on in-group favoritism postulates that respondents may be inclined to better accept discrimination stemming from the employer's biased point of view on the applicant's races (screening discrimination) [working hypothesis 3] .

2.2. Cost effects

Empirical and experimental evidence shows that the steepness of the trade-off between profit and morality is a determinant of discrimination in real-life markets. Zussman (2013) tested around 16,000 advertisements in the Israeli online market for used cars. Using a correspondence study, he showed that Arab buyers (whose ethnicity was suggested by their surname), received significantly less replies than Jew buyers did. He also found robust evidence of cost effects on discriminatory behaviors: the average gap between the replies to Arab buyers and the replies to Jew buyers disappeared when the Arab buyers offered a price 7.5% higher than the Jew buyers.

Tyran and Hedegaard (2018) showed in a recent field-experiment based on 169 secondary school Danish students that in-group favoritism exists and it is highly sensitive to the cost of discrimination. Tested students with Danish-sounding or Muslim-sounding names had to choose between working with a less productive student of his or her ethnic group or a more productive student of the other ethnic group. They found that 38% of the respondents preferred to discriminate, even if this decision was costly. However, the frequency of discrimination fell with the price of the discrimination: the probability of discrimination was

reduced by 9% when the cost of discrimination increased by 10%, resulting on an elasticity of 0.9 for the profit/morality trade-off⁷.

A public policy consequence of Zussman's (2013) and Tyran and Hedegaard's (2018) papers is that discriminatory behaviors can be affected by the steepness of the profit/morality trade-off, and that increasing the cost of discrimination through more efficient controls and sanctions could efficiently curb discriminatory behaviors.

In this paper, we are interested in knowing whether normative opinions about discriminatory behaviors can also be affected by this cost parameter. Dickson et al (2018) suggest such an effect, in the lab, for cases of favoritism; the aim here is to extend this result to different forms of discrimination. Our working hypothesis is that when the moral choice's cost is higher, respondents may be more inclined to condone the profitable but immoral behavior [working hypothesis 4].

2.3. Moral suasion effects

Last, being reminded that discriminatory behaviors are morally reprehensible may affect the respondent's opinion on their acceptability. Moral suasion effects (Romans, 1966) happen when an agent's behaviors and/or preferences are affected by a moral injunction, which can transit through different channels. For example, Dal Bó and Dal Bó (2014) found that observing an experimenter-crafted message with a moral standard affects both the expectations and the preferences of the participants to a voluntary contribution game. Tankard et al (2017) showed in a controlled experiment that the position emitted by an eminent institution, such as a US Supreme Court ruling supporting gay marriage, significantly modified the individuals' perceptions of norms. In this line, recent papers showed that leaders can trigger moral suasion by their actions (Mayer et al, 2013, Kesley and Recalde, 2015, Gächter and Renner, 2018) but also through statements to their followers. In an experimental setting, d'Adda et al. (2017) showed that unethical leaders have a causal influence on the ethical conduct of followers, through financial incentives but also by statements encouraging honest or dishonest behaviors.

In this paper, our goal is to check whether the profit/morality trade-off is sensitive to moral injunctions. We consider two kinds of moral injunctions: an egalitarian principle demanding an equal treatment of all groups and an affirmative action principle demanding a preferential treatment of the individuals who belonging to group who is discriminated against.

Our first objective is to confront respondents with moral injunctions that reflected real anti-discrimination public policies. Our working hypothesis is that these moral injunctions will have a significant impact on the respondent's normative preferences [working hypothesis 5].

Second, juxtaposing these two injunctions creates a quantitative gradation of the support asked in favor of the discriminated group. The point is to mimic the cost effect also tested in the protocol, and test the respondent's quantitative sensitivity to shocks that affect both sides of the morality/profit trade-off. Our protocol allows to compare the respondent's normative reaction to affirmative action to the more 'neutral' injunction of an equal treatment between races. Our working hypothesis is that respondents modulate their response to moral injunctions depending on the degree of support asked for the discriminated group. If this is true the respondents will be less forgiving of discrimination when faced by a moral injunction supporting affirmative action than by a moral injunction supporting an equal treatment of all candidates [working hypothesis 6].

Our last working hypothesis is based on a radically opposite premise. Various studies pointed out that quotas or positive discrimination can give the impression that discriminated groups receive more than their fair share of support and generate animosity against these groups (Valfort, 2018). In this case, we should find that affirmative action-based moral suasions are less effective in reducing the acceptability of discrimination than the more neutral equal treatment moral suasion [working hypothesis 7].

⁷ Using Zussman's (2013) online database, we observe that 18% of the Jew sellers who sent at least a reply (to a Jew buyer, an Arab buyer or both) prefer a Jew-buyer even if this buyer proposes a lower price. However, in this experiment the discrimination is also sensitive to cost effects: when the price gap between Jew and Arab buyers is lower than 500 US dollars, 16% of the Jew sellers choose an Arab buyer, but when the price gap is higher than 500 US dollars, the proportion rises to 27%.

3. Empirical strategy

To elicit causal motive, cost and moral suasion effects on the normative profit/morality trade-off, we use a quasi-experimental vignette protocol. This methodology is briefly presented in the next sub-section. The following subsections detail our protocol, the administration of the questionnaire and some descriptive statistics.

3.1. Vignettes in discrimination studies

Vignettes survey experiments are widely used in social sciences to study individual preferences and moral judgments⁸. The general principle is to present respondents with a dilemma randomly picked from a portfolio of short, controlled fictitious situations. Respondents are then asked to indicate their preferred outcome among a set of controlled alternatives that each reflects an alternative normative solution to the dilemma.

Papers on discrimination have so far used two kinds of vignette protocols (Zussman 2013, Baert and De Pauw, 2014, Finseraas et al., 2016). In the first kind of studies, groups of individuals are directly asked to express their feelings about a minority (Pager and Quillian, 2005, Chaiklin, 2011, Zussman 2013, Carrel et al., 2015). Respondents are presented with a very short non-contextualized direct question such as "Do you like members of group X?". The difference with simple attitudinal surveys is that the question is coupled either with a random assignment method or a testing. It is therefore possible to study the sensitivity of the answers to different treatments (for example the 'exposure' to a minority, such as in Carrel et al., 2015) or the correlation to actual discriminatory behaviors (Pager and Quillian, 2005, Zussman 2013).

In the second kind of studies, participants are presented with a story depicting a realistic situation (often a hiring decision) and asked to tell what they would do, in the real world, in this situation. Without the participants' knowledge, experimenters manipulate the fictional but realistic applicants' characteristics (often their gender or ethnicity). The fictitious hiring decisions are then compared across treatments, for example the exposure to a minority group (such as women in the armed forces, Finseraas et al., 2016), or the nature of discrimination (such as taste-based versus statistical discrimination, Baert and De Pauw, 2014).

Another kind of protocol is widely used in normative studies. In Empirical social choice papers, vignettes depict a controlled ethical dilemma along with a set of solutions that reflect conflicting normative principles. Each individual respondent is randomly assigned a story within a portfolio of parametric variations of the main vignette. By comparing the choices of the groups of respondents who received alternative scenarios, experimenters can derive causal effects of the manipulated parameters on normative preferences (see Atzmüller and Steiner, 2010, and Evans et al., 2015, for recent methodological surveys).

In this paper, we chose to follow the vignette protocol of normative studies. First, our objective was to explore discrimination from a normative point of view. We did not ask whether respondents like/trust members of groups targeted by discrimination (women, ethnic or religious minorities, immigrants), but want to provide evidence on how they solve the trade-off between the utilitarian principle (maximizing profit) and the moral principle (discrimination is wrong) at play in hiring decisions. A vignette depicting such a dilemma is a suitable way to achieve this goal. The realistic story-like format of vignettes reduces the cognitive effort of understanding the normative dilemma at play, and allows respondents to give their opinion without having to provide complex, overly long explanations (Alexander and Becker, 1978). Using stories set in fictitious settings also helps depersonalizing the issue by drawing the respondents away from their personal situation and social context (Schoenbert and Ravdal, 2000). Also, the information embedded in a vignette can be easily manipulated so that the respondent can be asked to balance a large set of conflicting factors and principles to make his or her choice (Wallander, 2009; Atzmüller and Steiner, 2010).

Second, focusing on normative preferences in a discrimination framework provides new insight on the relationship between attitudes and behaviors. Until recently, scholars reported major discrepancies between real-world and hypothetical, survey-collected behaviors (LaPiere, 1934, Pager and Quillian, 2005,

⁸ Survey experiments provide information on preferences, they are not designed to observe the actual behavior of the respondents.

Chaiklin, 2011), most studies concluding that attitudes are a bad proxy of actual behaviors. Recent research however shows that new evidence is needed to provide a full picture taking all motives of discrimination into account. When surveys are limited to questions such as ‘Do you like members of group X?’, the answers only provide information related to taste-based discrimination. If discriminatory real-world behaviors are mainly driven by screening or statistical motives, there is no reason why there should be any correlation between declared and observed behaviors. Zussman's (2013) paper on the discrimination in the Israeli second-hand car market substantiates this point. He used both a vignette protocol to collect attitudinal information on the prejudices of the buyers and a testing protocol to observe their real-world choices. He found no correlation between the Jewish buyer's declared opinions of the Arab sellers and their actions. However, he found that the opinion of the Arab buyers on the Jewish sellers' trustworthiness was significantly correlated with their actual choices, hinting at taste-based discrimination. In this case, vignette data did not predict discriminatory behaviors but provided information on the motive of discrimination. A similar reasoning can be found in Baert and De Pauw (2014).

Third, using vignettes allows the exploration of types of discrimination (such as customer taste or screening discrimination) that are difficult to re-create in an experimental setting. In the screening case, the key parameter is the decision-maker's intimate knowledge of the discriminated group; such a parameter is difficult to manipulate in the lab. In the client taste situation, a clever approach (as in Baert and de Pauw, 2014) is to ask participants whether they believe that their clients will be prejudiced against the discriminated group. This allows to study the propensity of the respondents to cater to their perceived customers and clients' prejudices. However, in this case discrimination originates in the beliefs of the experiment's decision-makers, who are real-world persons whose beliefs may be erroneous or fueled by their own (unobservable and maybe unconscious⁹) prejudices. Using vignettes, it is possible to declare that the decision-maker depicted in the scenario is absolutely free of prejudice and has certain information on the prejudices of his or her clients or collaborators. Vignettes allow experimenters to have a full control of the information available to the respondents. Scenarios can be as unambiguous, precise and complete as needed, so that vignettes provide flexible frameworks to test the whole range of the parameters at play in the theoretical mechanisms that are being explored.

Last, vignettes can lessen two of the main pitfalls of attitudinal surveys: reporting bias and social desirability bias. Asking directly sensitive questions about prejudice and discrimination could make the respondents too uncomfortable and/or suspicious of the experimenter's intentions (see Evans et al., 2015, for a general discussion of this issue). By contrast, vignettes' complexity lowers the likelihood of the respondent being fully aware of the factorial manipulation embedded in the vignette, thus reducing the social desirability bias (Wallander, 2009).

3.2. The survey

Between August 2018 and January 2019, 1,110 students from three French universities¹⁰ completed the survey. The anonymous, paper-and-pen survey was administered during lectures. It took about 20 minutes to complete. Students could opt-out from filling the survey and were not incentivized¹¹.

The survey had two parts: a vignette randomly drawn from a set of 15 alternative scenarios (see **Table 1**) and standard follow-up socio-demographic questions.

Our vignette was a short text depicting a hiring situation where an employer faced a profit/morality trade-off: either refusing to discriminate against an ethnicity and suffering a profit loss or choosing to discriminate and preserving his or her profit (see **Appendix B** for the exact phrasing of the vignette).

⁹ Implicit association tests developed in social psychology show that such discriminatory preferences can be unconscious (Greenwald et al., 1998), and that discriminatory behaviors may be influenced by contextual effects (Devine, 1989, Bertrand et al. 2005).

¹⁰ Paris-Descartes, Paris-Nanterre and Nouméa.

¹¹ This is standard practice in normative questionnaire-experiments. Since the point is to collect opinions on moral issues, incentivizing respondents would likely create a huge bias the results, *a fortiori* if one of the alternative principles tested is utilitarian and profit-based (see Gaertner and Schokkaert, 2011, for a methodological discussion).

To maximize the respondent's detachment with his or her own social context and personal experience, we set the story in a neutral and distant setting, a faraway planet where three equal-sized ethnic groups (humans and two alien races) lived in peace¹².

On this planet, a human restaurant owner needed to hire 10 new waiters among 20 candidates who belonged, in equal proportion, to 2 different races. We explicitly stated that the restaurant owner had no personal prejudice against any of the applicants' races. To induce a profit/morality trade-off and allow for a potentially justifiable discriminatory behavior on his or her part, we used factorial design to manipulate 3 parameters, resulting in a portfolio of 15 scenarios (see **Table 1**). In each scenario, 10 applicants among the 20 candidates belonged to a race that was discriminated against.

[Table 1 here]

Let's briefly present our factorial manipulations.

(1) The motive of discrimination:

- Customer taste discrimination (scenarios 1-5): the clients of the restaurant dislike being served by one group of applicants. The employer will lose clients and profit if he hires any waiter who belongs to this group ;
- Screening discrimination (scenarios 6-10): the employer does not know well the members of the discriminated group. He will lose profit if he hires any waiter from this race because he is not able to screen between good and bad waiters among members of this group;
- Statistical discrimination (scenarios 11-15): the employer has received information on the fact that the waiters from this group are on average less competent than others waiters. He might lose profit if he hires any waiter from this group. However, unlike in the screening discrimination case, there is no certainty that the employer will lose profit if he hires such candidate.

(2) The cost of morality, i.e. the % of profit loss incurred by the employer if he decides to hire any member of the targeted group:

- Low cost: 25% of the employer's clients and profit (scenarios 1, 6 and 11)
- Medium cost: 50% of the employer's clients and profit (scenarios 3, 4, 5, 7, 9, 10, 12, 14 and 15)
- High cost: 75% of the employer's clients and profit (scenarios 3, 8 and 13).

(3) Anti-discriminatory moral suasion effects, mediated by "*a leader whose authority is traditionally respected by all the members of the community*"¹³. This leader makes two kind of moral injunctions:

- Equalitarian norm (scenarios 4, 9 and 14): « *the respected leader said that one should hire an equal number of [members of the group that is discriminated against] and of [members of the other group]* »;
- Positive discrimination norm (scenarios 5, 10 and 15): « *the respected leader said that one should hire more [members of the group that is discriminated against] than [members of the other group] until further notice because of the hiring difficulties [the first group] face nowadays* »;
- A third case was the absence of moral suasion effects (scenarios 1-3, 6-8 and 11-13), where we introduced no information on a leader and provided no moral injunction.

Any given respondent was randomly assigned only 1 of the possible 15 versions of the vignette. This way, although he or she could immediately figure out that we were asking him or her about his or her opinion on discrimination, he or she was kept in the dark about the nature of the particular contextual factors that

¹² Having three equal-sized ethnic groups has two useful consequences. First, we are able to neutralize the employer's potential in-group favoritism: having three groups means that the employer can belong to a race not represented among the applicants, and have no intrinsic reason to prefer one race of applicants over the other. Second, equal-sized groups means that no race is demographically dominant, so we are able to leave out minority/majority and dominated/dominant issues.

¹³ Encompassing all the channels through which moral injunctions can transit (for example through vote-issued laws or generally accepted social norms) is beyond the scope of this paper. Using an universally respected leader, we bypass the need to specify a social choice procedure. Moreover it allows us to skip the issue of whether the scenario's protagonists got to vote for and/or agree with the law, and to establish that all of them (employers, applicants from the discriminated group and others) abide by the leader's decisions. Using a leader is also a handy way to avoid dealing with the plausibility of anti-discriminatory general social norms when our scenarios depict situations where a race is discriminated against.

were embedded in his or her version of the vignette. As a result, by comparing the answers from the 15 groups of respondents, it is possible to derive causal effects on the acceptability of discriminatory behavior of the 3 parameters manipulated in the protocol. Two extensions of the main questionnaire then presented to a subgroup of 110 additional respondents (see below in Section 4.4. for more details).

3.3. Variables and descriptive statistics

The respondents were asked to indicate, in their opinion and from a moral point of view, how many waiters belonging to the discriminated group it would be fair for the restaurant owner to hire¹⁴. This response variable could vary from 0 (no applicant from the target group should be hired) to 10 (all 10 vacant positions should only be filled by applicants from the target group). We interpreted this value as reflecting the intensity of discrimination's acceptability in the respondent's eyes.

If the respondent answered that about 5 members of the discriminated group should be hired, we interpreted his or her answer as not endorsing discrimination. If he or she however selected a value inferior to 4, we interpreted his or her choice as a statistically significant deviation from an equal consideration of both races of applicants, and as an endorsement of discriminatory behaviors. We distinguished between a weak (2 to 3 applicants should be chosen among the target group) and strong support of discrimination (only 1 to no applicant hired from this group). The resulting interest variable is therefore a categorical variable with three categories: (1) no support of discrimination, (2) weak and (3) strong acceptability of discrimination.

Follow-up questions were included to control for the homogeneity of the 15 groups of respondents. We collected information on the respondent's gender (male, female), university location, college major (economics, law, other major) and college year (first year, second or third year). The last two questions were attitudinal questions on discrimination. We asked the respondents whether unequal treatments based on gender, ethnicity or sexual orientation existed in their country and whether such unequal treatments were sometimes acceptable or always unfair (see **Table 2** for descriptive statistics).

[Table 2 here]

4. Results and discussion

This section is organized as follows. First, we present results on the general acceptability of discrimination. Second, we discuss how the profit/morality moral trade-off is affected by information on the motive and the consequences of discrimination. Third, we present results introducing moral suasion effects. The fourth subsection discusses two smaller-scale extensions of the protocol: an extension where humans (and not aliens) are the group targeted by discrimination and an extension where respondents are made aware of one of the treatments manipulated in the protocol.

4.1. Elicited vs attitudinal support of discrimination

Let's first consider the general acceptability of discrimination, using the data from the 9 scenarios devoid of moral injunctions (scenarios 1-2-3, 6-7-8 and 11-12-13 in **Table 1**, 556 respondents). We find that when asked directly about their attitudinal opinion on discrimination, 78% of our respondents (431 over 556) declare that unequal treatments based on gender, religion, origin or sexual orientation are unfair. Only 12% declare that differentiated treatments can be sometimes justified (67 over 556) and only 10% declare that such differentiated treatments do not exist in their country (58 over 556) (see results in **Table 3**).

[Table 3 here]

¹⁴ The question was not how many applicants the respondents would themselves hire if they were in the employer's place. Normative survey experiments provide information on preferences, and are not designed to ask about hypothetical behaviors. To avoid any confusion in the respondent's mind, we stated that "your opinion [as a respondent] on what is just will have no effect on Akri's actual hiring decision".

By contrast, in the vignette scenarios where we elicit a trade-off between discrimination and profit, we find that about 40% of the respondents believe that some labor market discrimination is fair (results in **Table 3**): 15% of the respondents strongly support discrimination, stating that it is fair to hire zero to one applicant from the discriminated group, and 25% of the respondents support a weaker form of discrimination, stating that it is fair to hire 2 or 3 applicants from the discriminated group.

Our first result is therefore that although respondents condemn discrimination, when confronted to a profit/morality trade-off, a large minority actually choose profit over morality.

Interestingly, we also find that there is no clear-cut correlation between the elicited and attitudinal acceptability of discrimination¹⁵. Scenario-elicited discrimination (either strong or weak) is acceptable for 39% of the respondents who stated that discrimination exists and is unfair when asked about their attitudinal opinion. Symmetrically, 51% of the respondents who declare that discrimination is sometimes justified do not find discrimination acceptable in the vignette. This means that the trade-off between profit and morality creates dilemmas that agents solve in ways that are not necessarily well predicted by their stated normative views. This finding echoes with previous findings on the link between attitudes and behaviors, several authors pointing at a very weak relationship between expressed opinions on discrimination and discriminatory behaviors (Pager and Quillian, 2005). Our data compounds this evidence on the normative acceptability of discriminatory behaviors.

4.2. Tipping the profit/morality trade-off: motive and cost effects

Next, we find that contextual information on both the motive of discrimination and the cost of morality has significant effects on the respondents' tolerance of discrimination.

Motive effects

We find that the acceptability of discrimination is different according to its causal mechanism (results in **Table 4**). When discrimination is customer taste-based (scenarios 1-3), 47% of the respondents declare that some discrimination is fair: 26% of the respondents declare that hiring 2 to 3 applicants from the disliked group is fair (weak discrimination), while 21% of the respondents declare that hiring zero to one disliked applicants is fair (strong discrimination). When the employer discriminates because of his or her poor ability to screen good applicants from a group he has little knowledge of (scenarios 6-8), a similar proportion of the respondents (43%) declare that some discrimination is acceptable. However, in this case the proportion of respondents who support a strong discriminatory behavior is only 13%. By contrast, in the statistical discrimination case, where the employer discriminates because he has heard that the workers from the target group are less productive (scenarios 11-13), only 29% of the respondents support discrimination.

[Table 4 here]

Statistical tests (see **table 4**) confirm that respondents are significantly more willing to discriminate when discrimination is justified by either customer preferences or screening than when it is justified by statistical discrimination. Such a hierarchy between the motives of discrimination echoes Baert and De Pauw (2014) who show that customer-based hiring discrimination is more prevalent than statistical discrimination¹⁶.

An interpretation may be that respondents may consider that the employer is a collateral victim of his or her customers' biased preferences, and be therefore more willing to exonerate his or her utilitarian behavior [working hypothesis 1]. Contrary to the two other cases, in customer taste discrimination, there is a distinction between the agents who are at the source of the discrimination and the agent who suffers

¹⁵ In all this paper, we use either the Kruskal-Wallis or the Jonckheere-Terpstra test to test the difference in the acceptability of discrimination. Note that several tests are available to test the relation between two qualitative variables or to test if a nominal outcome differs between k-groups: KHI-square test (Pearson, 1900), Kruskal-Wallis test (1954) and Jonckheere-Terpstra test (1952, 1954). The KHI-square test is recommended when both the outcome and the k-populations are non-ranked. The Kruskal-Wallis test is more powerful (higher probability that the test will reject the H0 when the H1 is true) when the outcome is ranked but not the populations. The Jonckheere-Terpstra test is more powerful when both the outcome and the populations are ordered. This *a priori* ordering of the k populations is due to the intensity of treatment (for more information, see Agresti et al., 1990).

¹⁶ The screening case is unfortunately not included in Baert and De Pauw's study (2014).

the financial consequences of the moral behavior. In addition, the adverse effect on the employer's profit is certain in the customer taste case, and only probable in the screening and statistical cases [working hypothesis 2].

However the screening and statistical discrimination cases are not treated in the same way by the respondents. An explanation could be the nature of the employer's misinformation. In the statistical discrimination case, the employer is in a situation of pure imperfect information: he has heard a rumor about a skill gap but has no information on the veracity of this rumor and the potential amplitude of the skill gap.

In the screening case however, there is an added factor: the employer comes from the same community than half of the applicants, and is uncertain about the characteristics of the applicants coming from another group. Respondents may find screening discrimination more acceptable than rumor and pure imperfect discrimination in the statistical discrimination case because they endorse the employer's rational homophilia or in-group favoritism [working hypothesis 3].

Cost effects

We also find a positive and significant relation between the cost of morality (i.e. a non-discriminatory behavior) and discrimination acceptability (**Table 5**) [working hypothesis 4].

The higher the cost of morality, the higher the proportion of respondents who support discrimination: when the cost of morality is the lowest (25% loss of clients and profit), only 33% of the respondents support some discrimination. Raising the cost of morality tips the profit/morality scale: when the cost is the highest (75% loss of clients and profit), 46% of the respondents agree with a discriminatory behavior. This difference is significant at 5%¹⁷.

Following Tyran and Hedegaard (2018), we determine the elasticity of the cost of morality. We find that a 1% increase in the cost associated with moral behavior reduces the probability of adopting such a behavior by 0.97 percentage points. This order of magnitude, although it is a discrete measure, is close to the one obtained by Tyran and Hedegaard (2018). Our results show evidence that by financially sanctioning discriminatory behavior, the legislator may affect both the behavior and the moral choices of economic agents.

[Table 5 here]

Further, we also find that the sensitivity of the respondents to the cost of non-discriminatory behaviors differs with the motive of discrimination (see **Table A1** in Appendix A). If the motive of the discrimination is driven by customer's preferences, when the cost of non-discrimination rises from 25% to 75%, the proportion of respondents who tolerate discrimination jumps from 32% to 59%. The p-value associated with the Jonckheere-Terspstra test is 0.003. By contrast, if discrimination is justified by screening issues or by statistical discrimination, the jump is smaller (from 40% to 44% and from 27% to 32%) and not significant at the usual levels¹⁸.

Robustness check

To check whether these results are robust to the respondents' observable characteristics, we estimate an ordered probit level on the intensity (strong, weak or nil) of the support of discrimination (estimated probabilities and the marginal effects of the treatments in **Table 6**).

[Table 6 here]

¹⁷Jonckheere-Terspstra test rejects the null hypothesis of no dependence between the variables with an error of 1.6%. We find no significant difference between the median profit loss (50%) and either of the two extreme losses (25% or 75%).

¹⁸ The p-value of the Jonckheere-Terspstra test is 0.35 if discrimination is justified by screening issues and 0.23 if it is justified by statistical discrimination.

The covariates include the different treatments manipulated in our protocol (motive of discrimination, cost of morality, and moral injunctions). Control variables include the respondent's university location, gender (male, female), and college major (economics, law or other). The regression results confirm that the acceptability of discrimination in the labor market is significantly determined by both motive and cost effects.

4.3. Moral suasion

Our results show that both of the moral injunctions tested have causal effects on the morality/profit trade-off. Keeping the cost effect constant at 50%, we focused on 9 scenarios (526 respondents) where we intersect a manipulation of a moral injunction (no moral injunction, equal treatment of all groups, affirmative action) and differentiated motives of discrimination (customer taste, screening, statistical) (see scenarios 2, 4, 5, 7, 9, 10, 12, 14, and 15 in **Table 1**).

We find that the existence of moral injunctions has a huge influence on the respondents' willingness to discriminate (**Table 7**) [working hypothesis 5]. Compared to the scenarios with no moral suasion, when an equalitarian injunction is stated the proportion of respondents who agree with some discrimination drops from 41% to 33%. When an affirmative action moral injunction is stated, the proportion falls to 29%. These differences are significant between the neutral scenarios and the scenarios with both kinds of moral injunctions are consistent across all motives of discrimination (Appendix **Table A2**). Moreover, it is interesting to note that unlike cost effects where respondents react to the intensity of the shock, on the 'moral' side of the moral/profit trade-off, an injunction of variable intensity does not significantly affect the respondents' preferences: there is no significant difference between the effects of the two injunctions. Respondents did not, either, display adverse effects to affirmative action policies [working hypothesis 6 and 7 not verified].

[Table 7 here]

Controlling for the respondent's observable characteristics, a probit ordered regression confirms that, all things being equal, the introduction of a moral injunction has a significant impact on the expression of discriminatory preferences (results in **Table 8**). We find that when a moral injunction that either promotes an equal treatment of both groups of applicants or a positive discrimination of the group facing difficulties on the labor market is introduced, the amount of respondents who support discrimination is divided by four.

[Table 8 here]

Appendix **Table A2** shows the causal effect of moral injunctions for each of the three motives of discrimination considered in the protocol. We find that here again the respondents set apart statistical discrimination from the other two motives of discrimination.

In the statistical discrimination sub-sample (scenarios 12, 14, 15), we find no effect of either kind of moral injunction on the acceptability of discrimination. By contrast, for the other two motives of discrimination, we find a strong effect of moral suasion, which significantly reduces the acceptability of discrimination.

This result suggests that reducing the acceptability of discriminatory behaviors could be achieved using non-coercive moral injunctions affirming and promoting egalitarian norms. The law expressivity literature suggests that the more these injunctions echo the respondents' ideological values, the more they affect their preferences and behaviors (Tankard and Levy Paluck, 2017). In this line of reasoning, Appendix **Table A3** presents moral injunction effects by declarative attitude towards discrimination. We find that the respondents who declared that discrimination exists and is not fair are more influenced by an egalitarian moral injunction than others. These respondents support much less often than others a highly discriminatory choice.

4.4. Extensions

We supplemented the initial questionnaire with two methodological extensions, to address identity and unobserved heterogeneity issues.

Addressing identity issues: discrimination of humans

In the initial version of the survey, the respondent's identity (human) matched the employer's. The aim of the first extension was to check whether respondents were sensitive to identity issues.

We therefore created three additional scenarios (scenarios 16, 17, 18, $n = 195$, see **Table 9**) where we swapped the identity of the employer and the target group. In the new scenarios, there was no moral injunction, an intermediate profit loss, and varying motives of discrimination. They were therefore identical to the scenarios 2, 7 and 12 ($n = 167$) of **Table 1**, except that in new scenarios the employer was now and alien, there were 10 applicants from another alien group and 10 human applicants, and the members of the discriminated group were humans (see **Table 9**). The phrasing of the new scenarios was otherwise identical to the initial scenarios.

[Table 9 here]

We find (results in **Table 10**) that shifting the identity of the employer and the applicants has a massive impact on the respondent's opinions on the acceptability of discrimination. The respondents are more likely to support some discrimination when the target group is human (55%) than when it is alien (33%). The gap is greatest when the discrimination is taste-based (69% versus 38%) but it remains positive or the other two motives of discrimination (45% versus 29% for screening and 46% versus 27% for statistical discrimination). Statistical tests indicate that these differences are all significant at the 5% or 10% threshold.

Contrary to recent papers (see for example Feld et al., 2016), we therefore find no evidence of in-group favoritism: the respondents were not more likely to condemn discrimination when the target group was human (new scenarios) than when it was alien (initial scenarios), quite the opposite.

[Table 10 here]

Two explanations could be provided for this surprising result. A first explanation relies on the fact that in the distant planet Neutra, humans can be perceived as foreigners, while the aliens could be viewed as the local inhabitants of the planet. In this case the respondents could rely on a first-come, first-served principle to solve the employer's ethical dilemma. They may find that local alien applicants should be favored over immigrant humans. Indeed, immigration as an excuse for discrimination has been extensively documented by real-world empirical studies (see for example Mayda, 2006 and Keita and Valette, 2019). In the 'immigrant' hypothesis, respondents should find more acceptable discrimination in the new scenarios.

A second explanation could be that the respondents might be freer from political correctness when they themselves belong to the target group. When the target group is human, like themselves, the respondents might feel freer to adopt a utilitarian view of the morality/profit trade-off. Conversely, when the target group is alien, the respondents might have more moral qualms about endorsing a discriminatory solution.

Addressing unobserved group heterogeneity: explicit treatment

In the initial survey, respondents are randomly assigned a scenario from our protocol's 15-scenario portfolio. Doing so prevents conformity bias, since respondents are not aware of the factors that are manipulated in their particular version of the survey. As a result, causal contextual effects can be derived from the results. However, a drawback is that unobserved individual heterogeneity is not controlled since different treatments are compared on different individuals. To address this potential issue, we presented a

sub-sample of 26 respondents with three successive scenarios (scenarios 1-3 and 6-8) where the cost of morality was gradually higher.

The results obtained (see **Table A4**) reveal two points. First, when respondents are aware of the manipulation, their response to treatments is significantly greater. The proportion of respondents who support discriminatory behaviors increases from 46% to 69% when the cost of non-discrimination suffered by the employer increases from 25% to 75%. By contrast, this difference was only 13 points on an unmatched sample (see **Table 5**). We also find that respondents are more likely to choose extreme choices (respectively 50% and 61% choose strong discrimination when the cost of non-discrimination reach 50% and 75% of the total income). We believe that such an hyper-reaction to the treatment (+45 points for the paired sample against 2 to 3 points for the initial version) probably captures two effects: a causal treatment effect and a strong social desirability bias. In any event this outcomes supports the use of a single scenario randomly chosen among alternative treatments.

5. Conclusion

Many attitudinal surveys (European Social Survey for Europe, CNCDH and CEREQ-Générations for France) reveal that although most people strongly disapprove of discrimination, a high percentage of people report having been discriminated against.

The economic literature provides explanations of this apparent paradox, explaining that discriminatory behaviors can be rational and durable when agents must arbitrate between profit and a moral behavior (i.e., refusing to discriminate). Recent papers use experimental methods (testing, field or lab experiment) to measure not only the extent of discriminatory behaviors but also to study the causal effect of key factors such as the cost of non-discriminatory behavior (Zussman, 2013; Tyran and Hedegaard, 2018) or measures promoting experienced diversity (Finseraas et al., 2016). These studies are instrumental to the identification of the parameters on which effective anti-discrimination public policies should be based.

This paper focused on a public policy lever that is crucial but seldom discussed in the literature: the moral aspect of discrimination. Our purpose in this paper was to investigate the parameters that affect the moral acceptability of the morality/profit trade-offs faced by decision-makers. To do so, we used a survey-experimental protocols based on vignettes that revealed causal effects on normative preferences. This methodology complements the methods classically used to study discrimination (Bertrand et al, 2005) and provides insights in configurations where implementing behavior-based experimental approaches is problematic (for example when discrimination caused by customer preferences or screening). Our protocol was specified to collect evidence on the sensitivity of the respondent's morality/profit trade-off to three key factors: the motive of discrimination, moral suasion effects and cost effects.

Our results show that when confronted to a profit/morality trade-off, a majority (60%) of the respondents stick to morality and do not discriminate. However, we also found that discrimination's acceptability was higher when it stemmed from consumer preferences or the decision-maker's social myopia than when it was based on a group's bad reputation. This means that for the respondents, imperfect information is not an acceptable motive of discrimination. In addition, we found that respondents' opinions were affected by shocks on both sides of the moral/profit trade-off: as expected, higher profit loss result in a greater acceptability of discrimination. At the same time, we found that respondents responded favorably to a moral injunction issued by a respected moral authority.

In terms of public policy, our results suggest that moral suasion-laden information campaigns could be effective in influencing agents facing morality/profit dilemmas. Moreover, since cost effects are shown to significantly affect the acceptability of discrimination in the eyes of the respondents, as proposed by (Zussman 2013 and Tyran and Hedegaard 2018), a system of financial incentives and sanctions could be designed to tip the profit/morality trade-off in favor of morality. Last, public policies should put special emphasis on customer discrimination which seems to be more accepted than other kinds of discrimination.

Several extensions of the paper can be considered. First, it might be interesting to compare our results with data collected on respondents belonging to other socio-demographic groups (business or human resources managers, public decision-makers, members of minority groups, people of older generations).

Second, the protocol itself could be extended in several directions. We have deliberately neutralized factors that could also affect moral/profit trade-off, such as the question of the possible minority and/or socially dominated status of the discriminated group. Relaxing this parameter might shed additional light on identity issues and help understand why respondents are more likely to accept discrimination when they themselves belong to the target group. It could also provide more differentiated results on the sensitivity of respondents to moral injunctions who demand quotas or affirmative action to help discriminated minority groups. It would also be interesting to see if the normative trade-offs are the same when considering other markets such as the housing market. Next, we showed that moral injunctions significantly affect the profit/morality trade-off when they are delivered by a respected authority figure. It could be interesting to test whether alternative vectors of moral injunctions (general social norm, legal rules, peer pressure) are as effective. In the same vein, equal treatment and affirmative action were the two moral injunctions presented to the respondents as proxies for active anti-discriminatory public policies. Vignette-based protocols could be used to explore the perceived normative effects of other kinds of public policies, such as nudges or monetary or penal sanctions.

Last but not least, this protocol, destined to collect normative preferences, could be associated with a more traditional experimental protocol in a field or lab experiment to explore if and how normative preferences affect the respondent's actual behaviors and/or votes.

6. References

- Agresti, A., Mehta C., Patel, N. (1990) Exact Inference for Contingency Tables with Ordered Categories, *Journal of the American Statistical Association*, 85(410): 453-458.
- Aigner, D.J., Cain, G.G. (1977) Statistical Theories of Discrimination in Labor Markets. *Industrial and Labor Relations Review*, 30(2): 175-187.
- Alexander, C.S., Becker, H.J., (1978) The use of vignettes in survey research. *Public Opinion Quarterly*, 42: 93-104.
- Arneson, J., (2006) What Is Wrongful Discrimination?. *San Diego Law Review*, 43(4): 775-808
- Arrow, K. J. (1973) The Theory of Discrimination. In O. Ashenfelter and A. Rees (Eds.), *Discrimination in Labor Markets*. Princeton, NJ: Princeton University Press: 3-42.
- Atzmüller C., Steiner, P. (2010) Experimental vignette studies in survey research. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences*, 6: 128-138.
- Baert, S., De Pauw, A. S. (2014) Is ethnic discrimination due to distaste or statistics? *Economic Letters*, 125(2): 270-273.
- Barr, A., Lane, T., Nosenzo D. (2018) On the social inappropriateness of discrimination. *Journal of Public Economics*, 164(C): 153-164.
- Becker, G. (1957, 1971, 2nd ed.) *The Economics of Discrimination*. Chicago, University of Chicago Press
- Bertrand, M., Chugh, D., Mullainathan, S. (2005). Implicit Discrimination. *American Economic Review*, 95 (2): 94-98.
- Bertrand, M., Duflo, E. (2017) Field experiments on discrimination. *Handbook of Economic Field Experiments*, Vol. 1. Elsevier: 309-393.
- Carrell, S.E., Hoekstra, M., West, J.E. (2015) The impact of intergroup contact on racial attitudes and revealed preferences. *Working Paper 20940 National Bureau of Economic Research*.
- Chassonnery-Zaïgouche, C. (2012) L'économie des discriminations peut-elle se passer d'une philosophie économique des discriminations ? *Documents de travail du Centre d'Economie de la Sorbonne n°1206*.
- CEREQ (2017) Quand l'école est finie. Premiers pas dans la vie active de la Génération 2013, V. Henrard et V. Ilardi (eds.), CEREQ.
- CNCDH (2017) Rapport de la Commission nationale consultative droits de l'Homme, Paris, La documentation française.
- Chaiklin, H. (2011) Attitudes, Behavior, and Social Practice. *Journal of Sociology & Social Welfare*, 38(1): 31-54.
- Cornell, B., Welch, I. (1996) Culture, information, and screening discrimination. *Journal of Political Economy*, 104: 542-571.
- d'Adda, G., Pavanini, N., Darai, D., Weber, R. A., (2017) Do leaders affect ethical conduct? *Journal of the European Economic Association*, 15(6): 1177-1213.
- Dal Bó, E., Dal Bó, P. (2014) Do the right thing: The effects of moral suasion on cooperation. *Journal of Public Economics*, 117: 28-38.
- Devine, P. G. (1989) Stereotypes and Prejudice: Their Automatic and Controlled Components. *Journal of Personality and Social Psychology*, 56(1): 5-18.
- Dickinson, D., Maslet, D., Peterle, E. (2018) Discrimination as favoritism: The private benefits and social costs of in-group favoritism in an experimental labor market. *European Economic Review*, 104: 220-236.
- ESS Round 4: European Social Survey Round 4 Data (2008). Data file edition 4.5. NSD - Norwegian Centre for Research Data, Norway – Data Archive and distributor of ESS data for ESS ERIC. doi:10.21338/NSD-ESS4-2008.
- Evans, S. C., Roberts, M. C., Keeley, J.A., Blossom, J. B., Amaro, C. M., Garcia A. M., Odar Stough C., Canter K. S., Robles, R., Reed, G. M. (2015) Vignette methodologies for studying clinicians' decision-making: Validity, utility, and application in ICD-11 field studies. *International Journal of Clinical and Health Psychology*, 15: 160-170.
- Feld, J., Salamanca, N., Hamermesh, D.S. (2016) Endophilia or exophobia: beyond discrimination. *Economic Journal*, 126: 1503-1527.
- Finseraas, H., Johnsen, Å A., Kotsadam, A., Torsvik, G. (2016) Exposure to female colleagues breaks the glass ceiling. Evidence from a combined vignette and field experiment. *European Economic Review*, 90: 363-374.
- Gächter, S., Renner, E. (2018) Leaders as role models and 'belief managers' in social dilemmas. *Journal of Economic Behavior and Organization*, 154: 321-334.
- Gaertner, W., Schokkaert, E. (2011) *Empirical Social Choice: Questionnaire-Experimental Studies on Distributive Justice*, Cambridge University Press, Cambridge.

- Greenwald, A., McGhee, D., & Schwartz, J. (1998) Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, 74(6): 1464-1480.
- Guryan, J., Charles, K.K. (2013) Taste-based or statistical discrimination: the economics of discrimination returns to its roots. *Economic Journal*, 123(572): 417-432.
- Keita, S., Valette, J. (2019) 'Natives' Attitudes and Immigrants' Unemployment Durations, *Demography*, 56(3): 1023-1050.
- Kesley, J. B., Recalde, M. P. (2015) Leadership and the Voluntary Provision of Public Goods: Field evidence from Bolivia. *Journal of Public Economics*, 122: 80-93.
- LaPiere, R. (1934) Attitudes vs Actions, *Social Forces*, 13: 230-237.
- Lane, T. (2016) Discrimination in the laboratory: A meta-analysis of economics experiments. *European Economic Review*, 90: 375-402.
- Lang, K., Lehmann, J-Y. (2012) Racial Discrimination in the Labor Market: Theory and Empirics. *Journal of Economic Literature*, 50 (4): 959-1006.
- Mayda, A.M. (2006) Who Is Against Immigration? A Cross-Country Investigation of Individual Attitudes toward Immigrants. *Review of Economics and Statistics*, 88 (3): 510-530.
- Mayer, D. M., Nurmohamed, S., Treviño, L. K., Shapiro, D. L., Schminke, M. (2013) Encouraging Employees to Report Unethical Conduct Internally: It Takes a Village. *Organizational Behaviour and Human Decision Processes*, 121: 89-103.
- Neumark, D. (2018) Experimental Research on Labor Market Discrimination. *Journal of Economic Literature*, 56(3): 799-866.
- Oh, S. J., Yinger, J. (2015) What Have We Learned from Paired Testing in Housing Markets?. in *Why Housing Agents Discriminate: The Measurement, Causes, and Consequences of Housing Discrimination*, John Yinger, e-book.
- Pager, D., Quillian, L. (2005) Walking the Talk? What Employers Say Versus What They Do. *American Sociological Review*, 70(3): 355-380.
- Phelps, E.S., (1972) The statistical theory of racism and sexism. *American Economic Review*, 62(4): 659-661.
- Pinkston, J. (2003) Screening discrimination and the determinants of wages. *Labor Economics*, 10(6): 643-658.
- Romans, J. T. (1966) Moral Suasion as an Instrument of Economic Policy. *American Economic Review*, 56(5): 1220-1226.
- Ross, S L., Yinger, J. (2002) *The Color of Credit: Mortgage Lending Discrimination, Research Methodology, and Fair-Lending Enforcement*, Cambridge, MA: MIT Press.
- Schoenberg, N., Ravdal, H. (2000) Using vignettes in awareness and attitudinal research. *International Journal of Social Research Methodology*, 3 (1): 63-74.
- Tankard, M. E., Levy Paluck, E. (2017) The effect of a supreme court decision regarding gay marriage on social norms and personal attitudes. *Psychological Science* 28(9): 1334-1344.
- Tyran, J-R., Hedegaard, M. (2018) Price of Prejudice. *American Economic Review. Applied Economics*, 10(1): 40-63.
- Valfort, M-A. (2018) Do anti-discrimination policies work?. *IZA World of Labor*, 450.
- Wallander, L. (2009) 25 years of factorial surveys in sociology. A review. *Social Science Research*, 38: 505-520.
- Welch, F. (1967) Labor-Market Discrimination: An Interpretation of Income Differences in the Rural South. *Journal of Political Economy*, 75(3): 225-240.
- Zussman, A. (2013) Ethnic Discrimination: Lessons from the Israeli Online Market for Used Cars. *Economic Journal*, 123(11): 433-468.

7. Tables

Table 1. Scenarios

Motive of discrimination	Moral injunction	Cost of non-discrimination (% of the employer's earnings)		
		25%	50%	75%
Customer taste	None	Scenario 1 ♥ N: 60	Scenario 2 ♥ ♣ N: 61	Scenario 3 ♥ N: 76
	Equal treatment		Scenario 4 ♣ N: 93	
	Positive discrimination		Scenario 5 ♣ N: 55	
Screening discrimination	None	Scenario 6 ♥ N: 70	Scenario 7 ♥ ♣ N: 53	Scenario 8 ♥ N: 59
	Equal treatment		Scenario 9 ♣ N: 55	
	Positive discrimination		Scenario 10 ♣ N: 68	
Statistical discrimination	None	Scenario 11 ♥ N: 57	Scenario 12 ♥ ♣ N: 53	Scenario 13 ♥ N: 67
	Equal treatment		Scenario 14 ♣ N: 41	
	Positive discrimination		Scenario 15 ♣ N: 47	

♥ Scenarios without moral suasion, used to study motive and cost effects ♣ Scenarios used to study moral suasion effects. Source: Cagou, 2019

Table 2. Descriptive statistics

	N	%
Socio-demographic characteristics		
<i>Gender</i>		
Male	338	37%
Female	577	63%
<i>College location</i>		
Nouméa	232	25%
Paris area	683	75%
<i>College year</i>		
First year	828	90%
2nd and 3rd year	87	10%
<i>College major</i>		
Law	399	44%
Economics	385	42%
Other major*	131	14%
<i>Standard of living</i>		
Low	287	31%
High	628	69%
Attitudes relative to unequal treatments		
<i>Unequal treatments exist</i>		
Yes	826	90%
No	89	10%
<i>Unequal treatment is sometimes fair conditional to the existence of the existence of unequal treatments</i>		
Yes	120	15%
No	706	85%
Number of observations	915	100%
* medical studies, management. Source: Cagou, 2019		

Table 3. Elicited and attitudinal acceptability of discrimination

Elicited acceptability of discrimination (number of applicants hired from the discriminated group)	All scenarios ♥		Attitudinal acceptability of discrimination					
			Some groups face unequal treatment				No group faces unequal treatment (2)	
			Yes and it is unfair (1)		Yes and it is fair (3)			
	N	%	N	%	N	%	N	%
Strong (0-1)	86	15%	65	15%	12	18%	9	16%
Weak (2-3)	137	25%	103	24%	21	31%	13	22%
None (4 and more)	333	60%	263	61%	34	51%	36	62%
Total	556	100%	431	100%	67	100%	58	100%
For (2), (3) and (1): Kruskal-Wallis Test (p-value) 44,553 (0.272)								
For (2) against (1) and (3): Kruskal-Wallis Test (p-value) 41,398 (0.133)								
♥ Scenarios without moral injunctions: scenarios 1-2-3, 6-7-8 and 11-12-13. Source: Cagou, 2019								

Table 4. Motive effects

Acceptability of discrimination (number of applicants hired from the discriminated group)	All scenarios ♥		Motive of discrimination					
			Customer taste discrimination		Screening discrimination		Statistical discrimination	
	N	%	N	%	N	%	N	%
Strong (0-1)	86	15%	42	21%	24	13%	20	11%
Weak (2-3)	137	25%	51	26%	54	30%	32	18%
None (4 and more)	333	60%	104	53%	104	57%	125	71%
All	556	100%	197	100%	182	100%	177	100%
Kruskal-Wallis Test (p-value) 13.9*** (0.001)								
Note: Kruskal-Wallis test show that respondents are significantly more willing to discriminate according the motive of discrimination (p-value = 0.001). We also tested solely customer taste against statistical and screening against statistical discrimination. The p-value are respectively 0.000 and 0.016								
♥ Scenarios without moral injunctions: scenarios 1-2-3, 6-7-8 and 11-12-13. Source: Cagou, 2019								

Table 5. Cost effects

Acceptability of discrimination (number of applicants hired from the discriminated group)	All scenarios ♥		Cost of morality (% of profit loss)					
			25%		50%		75%	
	N	%	N	%	N	%	N	%
Strong (0-1)	86	15%	24	13%	28	17%	34	17%
Weak (2-3)	137	25%	38	20%	40	24%	59	29%
None (4 and more)	333	60%	125	67%	99	59%	109	54%
All	556	100%	187	100%	167	100%	202	100%
Jonckheere-Terpstra (p-value): 38,664.5** (0.016)								
♥ Scenarios without moral injunctions: scenarios 1-2-3, 6-7-8 and 11-12-13. Source: Cagou, 2019								

Table 7. Moral suasion effects

Acceptability of discrimination (number of applicants hired from the discriminated group)	All scenarios ♥		Moral injunction					
			Equal treatment		Positive discrimination		No moral injunction	
	N	%	N	%	N	%	N	%
Strong (0-1)	63	12%	22	12%	13	8%	28	17%
Weak (2-3)	117	22%	40	21%	37	22%	40	24%
None (4 and more)	346	66%	127	67%	120	71%	99	59%
All	526	100%	189	100%	170	100%	167	100%
Kruskal-Wallis test (p-value)			2.73* (0.098)		6.09** (0.013)			
			5.64** (0.018)					
♣ Scenarios 2, 4, 5, 7, 9, 10, 12, 14, and 15. Source: Cagou, 2019.								
*** Significant at 1%, ** significant at 5%, * significant at 10%.								

Table 6: Estimated probabilities using an ordered probit model (scenarios without moral suasion, n= 556)

		Acceptability of discrimination					
		Strong		Weak		None	
All scenarios		14.4***	[11.4-17.4]	25,3***	[21.6-29.0]	60.3***	[56.2-64.4]
Cost of morality	25% vs 50%	NS		NS		NS	
	25% vs 75%	+6.9**	[+1.4; +12.3]	+4.8**	[+0.9; +8.7]	-11.7**	[-20.9; -2.5]
Motive of discrimination	Statistical vs taste-based	+6.1**	[+11.2; +0.9]	+5.1**	[+9.5; +0.7]	-11.2**	[-1.8; -20.6]
	Statistical vs screening	+10.0***	[+15.5; +4.5]	+7.3***	[+11.5; +3.1]	-17.3***	[-8.0; -26.6]
	Taste-based vs screening	ns		ns		ns	
Scenarios 1-3, 6-8, 11-13. *** Significant at the 1% threshold ** significant at the 5% threshold, *significant at the 10% threshold. NS: not significant. 95% confidence intervals are in square brackets. Covariates: location, gender and university major. Source: Cagou, 2019							

Table 8: Estimated probabilities using an ordered probit model (scenarios with moral suasion and 50% profit loss, n = 526)

		Acceptability of discrimination					
		Strong		Weak		None	
Panel B: Scenarios with moral suasion and 50% profit loss (n = 526 ♣)							
All scenarios		12.0***	[11.2-15.6]	22.2***	[20.8-26.2]	65.8***	[60.2-66.5]
Moral injunction	None vs equality	-6.2**	[-11,8; -0,5]	-4,8**	[-9,2; -0,4]	+11,0**	[6,9; 32,7]
	None vs affirmative action	-7,6***	[-13,1; -2,1]	-6,3***	[-10,9; -1,8]	+13,9***	[4,2; 23,6]
	Equality vs affirmative action	ns		ns		ns	
Scenarios 2, 4, 5, 7, 9, 10, 12, 14, and 15. *** Significant at the 1% threshold ** significant at the 5% threshold, *significant at the 10% threshold. NS: not significant. 95% confidence intervals are in square brackets. Covariates: location, gender and university major. Source: Cagou, 2019							

Table 9. Additional scenarios

Motive of discrimination	Identity of the employer and the target group	
	Human employer Alien target group	Alien employer Human target group
Customer taste	Scenario 2 N: 61	Scenario 16 ♦ N: 75
Screening	Scenario 7 N: 53	Scenario 17 ♦ N: 51
Statistical discrimination	Scenario 12 N: 53	Scenario 18 ♦ N: 69

♦ Additional scenarios. Source: Cagou, 2019

Table 10. Moral suasion effects on discrimination acceptability

Ethnic origin of the target group		Acceptability of discrimination (number of applicants hired from the discriminated group)						Total N
		Strong (0-1)		Weak (2-3)		None (4 and more)		
		N	%	N	%	N	%	
All	All	78	20%	91	24%	215	56%	384
	Human	56	29%	51	26%	88	45%	195
	Alien	22	12%	40	21%	127	67%	189
Customer taste	All	51	30%	36	21%	81	48%	168
	Human	37	49%	15	20%	23	31%	75
	Alien	14	15%	21	23%	58	62%	93
Screening	All	14	13%	25	24%	67	63%	106
	Human	9	18%	14	27%	28	55%	51
	Alien	5	9%	11	20%	39	71%	55
Statistical	All	13	12%	30	27%	67	61%	110
	Human	6	14%	13	32%	22	54%	41
	Alien	5	7%	14	20%	50	73%	69

♦ Scenarios 2, 7, 12, 16-18. Source: Cagou, 2019.

Appendix A

Table A1. Cross effects of cost and motive of discrimination

Acceptability of discrimination (number of applicants hired from the discriminated group)	Cost of morality						All scenarios ♥	
	25%		50%		75%			
	N	%	N	%	N	%	N	%
Customer taste								
Strong (0-1)	10	17%	12	20%	20	26%	42	21%
Weak (2-3)	9	15%	17	28%	25	33%	51	26%
None (4 and more)	41	68%	32	52%	31	41%	104	53%
All	60	100%	61	100%	76	100%	197	100%
Jonckheere-Terpstra (p-value): 4,764.0*** (0.003)								
Screening								
Strong (0-1)	9	13%	8	15%	7	12%	24	13%
Weak (2-3)	19	27%	16	30%	19	32%	54	30%
None (4 and more)	42	60%	29	55%	33	56%	104	57%
All	70	100%	53	100%	59	100%	182	100%
Jonckheere-Terpstra (p-value): 4,576(0.35)								
Statistical discrimination								
Strong (0-1)	5	9%	8	15%	7	10%	20	11%
Weak (2-3)	10	18%	7	13%	15	22%	32	18%
None (4 and more)	42	74%	38	72%	45	67%	125	71%
All	57	100%	53	100%	67	100%	177	100%
Jonckheere-Terpstra (p-value): 3,350 (0.23)								
♥ Scenarios 1-3, 6-8, 11-13. Source: Cagou, 2019.								

Table A2. Moral suasion effect on the number of persons hired from the discriminated group

Acceptability of discrimination (number of applicants hired from the discriminated group)	Moral injunction						All scenarios ♣	
	No moral injunction		Equal treatment		Positive discrimination			
	N	%	N	%	N	%	N	%
Customer preferences								
Strong (0-1)	12	20%	14	15%	4	7%	30	14%
Weak (2-3)	17	28%	21	23%	10	18%	48	23%
None (4 and more)	32	52%	58	62%	41	75%	131	63%
All	61	100%	93	100%	55	100%	209	100%
Kruskal-Wallis Test (p-value)			1.45 (0.228)		6.53** (0.011)			
			4.01** (0.045)					
Screening								
Strong (0-1)	8	15%	5	9%	6	9%	19	11%
Weak (2-3)	16	30%	11	20%	18	26%	45	26%
None (4 and more)	29	55%	39	71%	44	65%	112	64%
All	53	100%	55	100%	68	100%	176	100%
Kruskal-Wallis Test (p-value)			2.97* (0.085)		1.5 (0.221)			
			2.83* (0.0925)					
Statistical discrimination								
Strong (0-1)	8	15%	3	7%	3	6%	14	10%
Weak (2-3)	7	13%	8	20%	9	19%	24	17%
None (4 and more)	38	72%	30	73%	35	74%	103	73%
All	53	100%	41	100%	47	100%	141	100%
Kruskal-Wallis Test (p-value)			0.13 (0.712)		0.28 (0.594)			
			0.28 (0.59)					
♣ Scenarios 2, 4, 5, 7, 9, 10, 12, 14, and 15. Source: Cagou, 2019.								

Table A3. Attitudinal opinion on discrimination and moral suasion

Opinion on discrimination	Moral suasion effect	Acceptability of discrimination (number of applicants hired from the discriminated group)					
		Strong (0-1)		Weak (2-3)		None (4 and more)	
		N	%	N	%	N	%
Fair or does not exist	None	5	12%	15	38%	20	50%
	Equal or positive	11	13%	21	25%	52	62%
	Kruskal-Wallis Test (p-value): 1.062 (0.3028)						
Unfair	None	23	18%	25	20%	79	62%
	Equal or positive	25	9%	55	20%	195	71%
	Kruskal-Wallis Test (p-value): 4.398** (0.036)						

Table A4: acceptability of discrimination in a paired sample

Acceptability of discrimination (number of applicants hired from the discriminated group)	Cost of morality						Total	
	25%		50%		75%			
	N	%	N	%	N	%	N	%
Strong (0-1)	4	15%	13	50%	16	61%	33	42%
Weak (2-3)	8	31%	5	19%	2	8%	15	19%
None (4 and more)	14	54%	8	31%	8	31%	30	39%
Total	26	100%	26	100%	26	100%	78	100%
Jonckheere-Terpstra test (p-value): 8.92** (0.012). Source: Cagou, 2019.								

Appendix B. Full text of the vignette

All the versions of the vignette (see below, translated from French) were structured the same way.

A short introduction, common to all scenarios, presented the general setting and the hiring decision.

Next followed

- information on the source of discrimination (1 variant randomly assigned among the 3),
- information on the cost of the moral non-discriminatory alternative (1 variant randomly assigned among the 3, worded so as to keep the story consistent with the information given about the source of information),
- information on moral suasion (1 variant randomly assigned among the 3).

All versions of the vignette ended with the same closing statement and the same question.

GENERAL SETTING: COMMON TO ALL SCENARIOS

Planet Neutra is a faraway planet located in a peaceful galaxy. In this planet, the population is composed by humans and by two kinds of aliens, the Tenkas and the Tokrins. There is the same number of humans, of Tokrins and of Tenkas on the planet.

Akri is the human manager of a restaurant located on Neutra. For the new season, he must hire 10 new waiters. Personally, Akri equally likes all the human and alien races, and believes that they are equally good waiters.

INFORMATION ON THE SOURCE OF DISCRIMINATION: 3 VARIANTS

[Customer taste] Akri has noticed that customers do not like to be attended by Tenka waiters. There are 20 applicants: 10 Tenkas and 10 Tokrins.

[Screening] Akri has always lived among humans and has not had many occasions to mingle with aliens in this personal or professional life. As a result, he doesn't know well the alien races and has poor insight on their education, their habits, their punctuality, and their behavior with the customers. He knows that there are good and bad waiters among all races. Because he is human, Akri knows that he is able to tell good human waiters from bad human waiters, and hire only good waiters among human applicants. However, he is not able to tell the difference between good and bad *alien* waiters, so that over 10 alien applicants he will hire 5 bad waiters. There are 20 applicants: 10 Tenkas and 10 humans.

[Statistical] Akri has always lived among humans and has not had many personal or professional occasions to mingle with aliens. As a result, he doesn't know well the alien races and has poor insight on their education, their habits, their punctuality, and their behavior with the customers. However, he knows that there are good and bad waiters among all races. The good waiters are liked by the customers and the bad waiters drive the restaurant's customers away. Akri has heard that there are more bad waiters among the Tenkas than among the Tokrins. He does not know whether this information is true or false, and does not have enough information to tell the difference during the hiring process. It is important for Akri to hire good waiters because the restaurant is his only source of income. There are 20 applicants: 10 Tenkas and 10 Tokrins.

INFORMATION ON THE COST OF MORALITY: 3 VARIANTS

[Cost low] The restaurant loses a quarter of his customers if there are Tenka waiters [if Source is Customer taste] / bad waiters [if Source is Screening or Statistical]. The restaurant loses no clients if there are only Tokrin waiters [if Source is Customer taste] / good waiters [if Source is Screening or Statistical].

[Cost medium] The restaurant half of his customers if there are Tenka waiters [if Source is Customer taste] / bad waiters [if Source is Screening or Statistical]. The restaurant loses no clients if there are only Tokrin waiters [if Source is Customer taste] / good waiters [if Source is Screening or Statistical].

[Cost high] The restaurant three quarter of his customers if there are Tenka waiters [if Source is Customer taste] / bad waiters [if Source is Screening or Statistical]. The restaurant loses no clients if there are only Tokrin waiters [if Source is Customer taste] / good waiters [if Source is Screening or Statistical].

INFORMATION ON MORAL SUASION: 3 VARIANTS

[Moral none] --- no text included ---

[Moral affirmative action] There is on Planet Neutra a great leader whose authority is traditionally respected by all members of the community (humans, Tokrins and Tenkas). This great leader said that only Tenkas should be hired until further notice because of the hiring difficulties they face on the labor market.

[Moral equal treatment] There is on Planet Neutra a great leader whose authority is traditionally respected by all members of the community (humans, Tokrins and Tenkas). This great leader said that an equal number of Tokrins (if Sources 1 and 3) / Tokrins (if source 2) and of Tenkas should be hired.

CLOSING STATEMENT AND QUESTION: COMMON TO ALL SCENARIOS

Having a profitable restaurant is important for Akri because it is his only source of income.

In your opinion, how many Tenka waiters it fair for Akri to hire? Your opinion will have no effect on Arki's actual decision.

0 - 1 - 2 - 3 - 4 - 5 - 6 - 7 - 8 - 9 - 10