Hiring in a Diverse Labor Market: Spillovers Across Minority Groups*

Daniel Gómez-Vásquez†

December 2025

Click here for the latest version.

Abstract

I study whether hiring experiences with workers from one minority group impact beliefs about performance and hiring of workers from a different group. I conduct an online hiring experiment with a geographically representative U.S. sample, where White employers decide whether to hire White, Black, or Hispanic workers, in a setting where there are no differences in performance across groups. In the experiment, guided by a Bayesian framework featuring asymmetric updating on negative versus positive experiences, I manipulate the worker groups available for hire. I employ the experiment first to identify and compare differences in beliefs about performance and hiring against Black and Hispanic workers. I then investigate whether hiring experiences with Black (Hispanic) workers have positive or negative spillovers on beliefs about performance and the hiring of Hispanic (Black) workers. Results from the experiment show that Black and Hispanic workers are believed to perform worse than White workers and are hired less often. I find positive hiring spillovers from Hispanic to Black workers only: while employers are more likely to hire Black workers after hiring experiences with Hispanic workers, the converse is not true. I provide suggestive evidence that the latter finding is driven by a more pronounced and persistent bias in beliefs against Hispanic workers.

Keywords: Hiring experiment; discrimination; belief updating; hiring spillovers.

JEL Codes: C91, D90, J15, J71, M51

*I am grateful for all the comments, support, and guidance from Danila Serra. I am indebted to Silvana Krasteva for her invaluable help and guidance with the conceptual framework. I thank Alan Benson, Benjamin Bushong, Marco Castillo, Alex Chan, Emily Cook, Stefano DellaVigna, Catherine Eckel, Jorge Garcia, Seth Gershenson, Peter Hull, Alex Imas, Joanna Lahey, Louis-Pierre Lepage, Marissa Lepper, Danyao Li, Ulrike Malmendier, Jonathan Meer, Katherine Milkman, Marco Palma, Ragan Petrie, Tim Salmon, Mallika Thomas, Lise Vesterlund, Maria Zhu, Anna Ziff, and participants in the FAIR NHH-UCSD Rady Spring School in Behavioral Economics and the Russell Sage Summer Institute in Behavioral Economics for helpful comments. Financial support from the Texas A&M Department of Economics and the College of Liberal Arts (now Arts & Sciences) is gratefully acknowledged. This work has also been supported by Grant #2302-42091 from the Russell Sage Foundation. Any opinions expressed are those of the principal investigator alone and should not be construed as representing the opinions of the Foundation. This study obtained ethical approval from the Texas A&M University Institutional Review Board (IRB2022-0331M). The online experiment is pre-registered on AsPredicted (#107386). All errors are my own.

†College of the Holy Cross, Department of Economics & Accounting, dgomez-vasquez@holycross.edu

1 Introduction

The United States is a multi-racial and multi-ethnic society, as reflected in the labor market. Blacks and Hispanics constitute more than one-third of the United States labor force, and unemployment rates are significantly higher for these minority groups than for White Americans.[1] In such a diverse society, employers likely make hiring decisions involving applicants from multiple minority groups, and experiences with workers from one minority group may affect beliefs and attitudes towards workers from a different minority group. Migration inflows, among other shocks that increase the availability of a traditionally disadvantaged minority group (hence their hiring), may in turn impact - positively or negatively - the hiring of workers from a different disadvantaged group. While hiring discrimination has been widely studied and documented in economics, showing evidence of discrimination against disadvantaged minority workers (e.g., Bertrand and Duflo, 2017; Neumark, 2018; Lang and Spitzer, 2020), the existing body of work on labor market discrimination has typically focused on hiring decisions and biases towards only one minority group as compared to the majority group. Studies involving multiple minority groups and investigations of spillovers across such groups are missing.

In this paper, I test for the existence, direction, and magnitude of bias in beliefs, hiring, and spillover effects across multiple minority groups. First, I ask whether there are biases in employers' beliefs about the performance of Black and Hispanic workers, and whether there is hiring discrimination. I then ask whether hiring experiences with workers from one minority group (e.g., Black) affect subsequent beliefs about performance and the hiring of a different minority group (e.g., Hispanic). I also explore whether biases in beliefs, hiring, and spillover effects differ by minority group.

Guided by a conceptual framework of hiring spillovers in the workplace, I design and implement a set of online hiring experiments involving White employers and White, Black, and Hispanic workers. I recruited a first group of participants to act as workers and perform a series of math algebraic problems.[2] This task is such that there are no statistically significant differences in average performances across the groups.[3] I then recruit a second group of participants to act as employers. Employers make two sets of 10 hiring decisions each, i.e., 20 hiring decisions in total. For each decision, they have to choose whom to hire between a randomly

[1] See the US Bureau of Labor Statistics 2023 and 2025 reports.

[2] I focus on an environment where White workers are believed to have higher performance than Black and Hispanic workers. Future work can extend this analysis to other racial or ethnic minorities. However, depending on the task, some minority groups might not be considered disadvantaged. For instance, Asian students tend to have higher scores than White students in the math section of the SAT (see the National Center for Education Statistics, 2023). Hence, one would expect more hiring of Asian than White workers for this task.

[3] I chose to employ a real-effort task without differences in average performance across ethnic and racial groups to abstract from the role that statistical differences between the groups may play in discrimination. Instead, I focus on the role that hiring experiences play in spillovers across groups.

chosen White worker and a randomly chosen Black or Hispanic worker, and get paid based on the hired worker's performance. While worker performance is unobservable before the worker is hired, employers learn the performance of the chosen worker after hiring, and a new hiring decision follows in the next period.

In order to study whether and how hiring experiences with one minority group may "spill over" to a different group, I implement four experimental treatments, which differ in the groups of workers an employer could hire from. In the HH treatment, employers can only hire workers from a group of White workers and a group of Hispanic workers. This means that, in each period of the two sets of 10 hiring decisions, the employer chooses which group she wants to hire from and is randomly assigned a worker from that group. The BB treatment is identical to the HH treatment, except that the employer now chooses from the group of White workers and the group of Black workers in both sets of hiring decisions. I call the treatments HH and BB the "baseline" treatments. The "spillover" treatments, BH and HB, differ in that employers, in the first set of 10 hiring decisions, can hire from White workers and from the set of workers belonging to one minority workers (e.g., Blacks), and, in the second set of 10 decisions, from White workers and from the other minority group (e.g., Hispanics). Specifically, in the HB treatment, employers choose between White and Hispanic workers in the first set of hiring decisions and choose between White and Black workers in the second set. In the BH treatment, employers choose between White and Black workers in the first set and choose between White and Hispanic workers in the second set. The baseline treatments serve as controls for the spillover treatments as they show the levels of bias in beliefs and hiring against Hispanic and Black workers, respectively, in a setting in which only workers from one minority are available for hire.

I disclose the racial or ethnic groups in a subtle way to limit social desirability bias. Specifically, I tell employers that the previous participants in the role of workers had been assigned to three different groups based on their demographics, such that Group Purple is made of White workers, whereas Groups Orange and Green are made of Hispanic/Latino or Black/African American workers. This information is given in the experiment instructions, and participants also respond to incentivized comprehension questions to make sure they are aware of the groups' composition. Hence, when making hiring decisions, employers choose between a "worker from Group Purple" and a "worker from Group Orange" instead of a "White worker" and a "Hispanic worker." To control for color preferences, the minority group that employers encounter in the first set of hiring decisions is always associated with Group Orange, and the minority group in the second set of the spillovers treatments is always associated with Group Green.

In order to measure initial bias against minority workers, I elicit employers' beliefs about the average performance of workers in Group Purple, Group Orange, and Group Green before the hiring task. Specifically, I tell employers the average performance, putting together all the

groups and workers, and then ask them the expected performance from the workers in each group.[4] Employers earn additional payment based on the accuracy of their beliefs and get feedback at the end of the experiment. I measure beliefs again in the middle of the hiring task (between the first and second set of hiring decisions) and at the end (after the second set). The initial measurement of beliefs before the hiring task captures the views, prejudice, or stereotypes that employers hold about the different groups, which could guide their initial hiring decisions. Tracking employers' beliefs over time allows me to assess the extent to which employers update their beliefs based on hiring experiences and whether experiences with workers from one minority group affect belief-updating about the performance of a different group.

The experiment generates three main findings. My first finding is that Black and Hispanic workers are hired less often than White workers, and their performance is underestimated, with hiring experiences affecting beliefs about performance and subsequent hiring. Specifically, before any hiring, employers believe that Hispanic workers performed 19% worse than White workers and that Black workers performed 18% worse than White workers in the math task. This is driven by a large mass of employers underestimating minority workers' performance (45% of the employers) rather than a few employers with extreme beliefs. Hispanic and Black workers are 17.6 and 14.1 percentage points less likely to be hired than White workers, respectively. Moreover, hiring experiences cause employers to update their beliefs about minority workers' performance: their posterior belief is a weighted average of their prior belief and the observed performance of the hired workers. Employers who have positive hiring experiences, i.e., hired workers with performance above their prior, increase their posterior beliefs, while there is a decrease for those who had negative hiring experiences. Experiences also matter for subsequent hiring, as a positive hiring experience increases the number of minority workers hired in the following periods. The opposite holds for negative hiring experiences.

My second finding is that employers are more biased against Hispanic workers: they have lower expectations about Hispanic workers' performance, and the hiring gap against Hispanic workers is larger. Moreover, there are striking differences in the belief updating processes regarding the performance of Hispanic and Black workers. Specifically, I find that employers tend to place more weight on prior beliefs over time, suggesting that first impressions with one group matter. This is particularly true for updating about Hispanic workers' performance. In addition, I find evidence of asymmetric updating only about Hispanic workers' performance, meaning that, when updating their beliefs, employers place (82 percent) more weight on negative than on positive experiences. This indicates that employers are pessimistic about Hispanic workers' performance. This is consistent with confirmation bias: employers expect Hispanic

[4]Telling employers the average performance of all the workers together allows me to minimize the chances of employers giving arbitrary numbers or to give extremely high numbers for White workers and low numbers to minority workers.

workers to perform badly, and they interpret information from hiring experiences in a way that reinforces their prior.

My third finding is that there are spillovers in beliefs and hiring across minorities. To begin, positive experiences with Black workers improve beliefs about Hispanic workers' performance. The opposite holds for negative experiences. However, this updating is asymmetric: only positive experiences with Hispanic workers impact beliefs about Black workers' performance. Put differently, negative experiences with Hispanic workers are not used by employers to update their beliefs about Black workers. Spillovers in hiring are also asymmetric. Specifically, there exist positive hiring spillovers from Hispanic to Black workers, but not from Black to Hispanic workers. Hiring experiences with Hispanic workers increase the likelihood that Black workers are hired, up to the point that employers are indifferent between hiring White and Black workers, but the opposite is not true; Hispanic hiring is unaffected by previous experiences with Black workers. This asymmetry is consistent with more bias against Hispanic workers than against Black workers. In fact, if employers see Black workers as more productive than Hispanic workers, (positive and negative) hiring experiences with Black workers affect (positively and negatively) beliefs about Hispanic workers' performance and, thus, their hiring once they become available for hire. Alternatively, positive experiences with Hispanic workers (the lowest-ranked group) may induce employers to update their beliefs about Black workers upward, also making them more likely to hire Black workers once they become available for hire. Negative experiences with Hispanic workers, on the other hand, may not be deemed informative and, instead, confirm the stronger bias against Hispanic workers while not affecting beliefs about Black workers' performance. Hence, these employers are not less likely to hire Black workers, as opposed to those whose previous negative experiences were with Black workers, and tend to hire fewer Black workers later on.

My exploratory analysis of mechanisms confirms this interpretation of the results. I first use my data on beliefs about performance and hiring to decompose the sources of discrimination into prejudice (or taste-based discrimination) and inaccurate statistical discrimination. While the decomposition exercise confirms the existence of substantial prejudice against both Black and Hispanic workers, prejudice against Hispanic workers is significantly larger. I then consider the role that natural exposure to Black and Hispanic populations may play in driving my findings. I exploit variation in the percentage of Black and Hispanic populations in the participants' counties of residence. This exploratory analysis generates insights that are in line with my experimental findings. First, natural exposure to Black individuals does not impact the hiring of Hispanic workers. Second, natural exposure to Hispanic individuals increases the likelihood that Black workers are hired. Third, greater natural exposure to Hispanic individuals diminishes the stickiness of beliefs about Hispanic workers' performance, i.e., increases the weight placed on experiences and decreases the weight placed on priors when forming posterior beliefs. Hence,

biases against Hispanic workers are more pronounced, and experiences with them, either in the hiring experiment or through natural exposure, make employers more likely to hire Black workers when available, since they are "better" ranked than Hispanics.

My findings corroborate a large body of evidence of bias against Black workers (Bertrand and Duflo, 2017). I find that employers believe that Black (Hispanic) workers performed 18% (19%) worse than White workers, and Black (Hispanic) workers are 14.1 (17.6) percentage points less likely to be hired than White workers, representing a 24.7% (30%) hiring gap. These estimates are comparable to those from Evsyukova® et al. (2025), who find that Black profiles on LinkedIn are 13% less likely to be accepted, and those from Quillian et al. (2017)'s meta-analysis, which concludes that White applicants tend to receive 36% more callbacks than Black applicants and 24% more callbacks than Latino applicants. However, the greater bias against Hispanic workers I find is not consistent with such evidence. One potential explanation is the very limited existing studies including Hispanic/Latino applicants (Quillian et al., 2017). Moreover, in the past years, political campaigns have contained racist and anti-immigrant, in particular, anti-Hispanic rhetorics, which might have also changed discrimination patterns, suggesting more bias against Hispanic workers, and thus explaining the increase in Black hiring following hiring experiences with Hispanic workers. Nevertheless, my findings are in line with those of Fouka and Tabellini (2022), as they show that Mexican immigration improved Whites' attitudes and behavior towards Black individuals, confirming a more severe bias against Hispanic workers and the existence of asymmetric spillover effects in hiring. Specifically, negative hiring experiences with Hispanic workers, the more recent and more socially distant minority, are not seen as informative and, therefore, not accounted for when making hiring decisions involving Black workers, the "resident" minority.

My paper contributes to the literature on how experiences or events involving one minority group may affect attitudes toward different minority groups. Besides Fouka and Tabellini (2022)'s study on how Mexican immigration improved outcomes for Black individuals, following the re-categorization of Black individuals as "Americans" since the inflows of a new (more distant) group redefine the attributes considered relevant for social categorization, a few studies have looked at how the influx of new minority groups into U.S. cities has shifted attitudes towards existing minority groups. For instance, Fouka et al. (2022) presents evidence that, during the First Great Migration, the arrival of Black individuals to the Northern cities facilitated the integration of formerly discriminated European immigrants into the White American Society. Other studies have explored how adverse events involving one minority group may worsen attitudes toward other minority groups. For instance, McConnell and Rasul (2021) show how Hispanic defendants had worse judicial outcomes following the 9-11 attacks.[5] I add

[5]In the context of education, Shi and Zhu (2023) show that the presence of Asian students increases Black-White and Hispanic-White teacher assessment gaps in elementary and secondary schools.

to this literature by studying spillovers across minority groups in a labor market setting, and by providing further evidence of how hiring experiences with Hispanic (Black) workers may or may not re-shape beliefs and behaviors of White employers towards Black (Hispanic) workers.

My paper also contributes to the vast literature on labor discrimination. Labor market discrimination has been studied for over 60 years, following the classical views of taste-based (Becker, 1957) and statistical discrimination (Phelps, 1972; Arrow et al., 1973; Aigner and Cain, 1977).[6] Recent research has proposed other drivers of discrimination, such as motivated reasoning (Eyting, 2022), employers holding inaccurate beliefs (Bohren et al., 2025), or not acquiring information about candidates from less attractive groups (Bartoš et al., 2016). Studies more closely related to my paper investigate the role of personal experiences on hiring decisions (Benson and Lepage, 2024; Lepage, 2024; Leung, 2018).[7] I contribute to this literature by studying whether and how inaccurate beliefs regarding a disadvantaged group, and therefore the hiring of workers from that group, may be affected by personal experiences with workers from a different disadvantaged minority group. Hence, this paper adds to the broader literature on discrimination by studying multiple disadvantaged minority groups within the same labor market environment. Most of the existing studies have studied dynamics between two groups only, focusing on one majority and one disadvantaged minority group, or even considering a consolidated minority consisting of multiple groups.[8] Notable exceptions are the study by Chan (2022) on discrimination against Black and Asian doctors as opposed to White doctors, and Aaronson and Phelan (2022)'s investigation of job losses of White, Asian, and non-Asian workers. While these studies involve multiple minority groups, they show outcomes for each minority independently and do not focus on how the outcomes for one minority group could affect outcomes for another minority group.

Finally, my paper also adds to the literature on labor market discrimination against Black and Hispanic workers in the United States. The vast majority of these studies have shown evidence of White-Black hiring and earnings gaps in the labor market (Altonji and Pierret, 2001; Bertrand and Mullainathan, 2004; Benson and Lepage, 2024; Evsyukova® et al., 2025; Haaland and Roth, 2023; Hurst et al., 2024; Kline et al., 2024). A few audit and correspondence studies have focused on biases against Hispanic workers, failing to find evidence of discrimination, with the exception of the early studies of Bendick Jr et al. (1991), which focused on Latinos

[6] Taste-based discrimination describes employers who dislike hiring members from a particular group, while statistical discrimination explains discrimination as a generalization of information about true group differences, leading to differential treatment.

[7] Other recent studies examine how employers may neglect signals and therefore discriminate more against disadvantaged groups (Campos-Mercade and Mengel, 2024), or how employers interpret information to be more representative than it really is when groups are presented together, generating discriminatory gaps and thus amplifying underlying group differences (Esponda et al., 2023).

[8] While Szymanski (2000), Ondrich et al. (2003) and Wozniak (2015) compare Blacks and non-Blacks, most of the studies consolidate the minorities as non-Whites, see List (2004, 2006); Burgess and Greaves (2013); Alesina and La Ferrara (2014); Goncalves and Mello (2021); and Kessler et al. (2024).

in Washington DC, and of Pager et al. (2009), which focused on Puerto Ricans in New York City.[9] I contribute to this literature by examining and directly comparing, within the same experimental labor market, White employers' biases in beliefs and hiring towards Black and Hispanic workers, as opposed to White workers, and by showing that in my setting, the bias is more pronounced and stickier for Hispanic than for Black workers.

The rest of the document is organized as follows: Section 2 presents the conceptual framework. Section 3 introduces the experiment design. Section 4 describes the estimation strategy. Section 5 presents the results. Finally, Section 6 contains the conclusion.

2 Conceptual Framework

In this section, I model spillovers of hiring experiences in a diverse labor market. Employers choose a worker to hire in a given time period. Workers who can be hired belong either to the majority – White – or to one of two disadvantaged minorities, Black or Hispanic. Spillovers across minority groups occur when employers' experiences with workers from one minority (e.g., Black) affect their perception and hiring of workers from another minority (e.g., Hispanic). I use this model to conceptualize how an employer may change their beliefs about performance of workers from a disadvantaged minority group following hiring experiences with workers from a different group and how these changes in beliefs affect hiring decisions.

I first introduce a baseline model of employers' beliefs and choices, highlighting the role of positive and negative hiring experiences in belief updating. I then show that employers could be pessimistic about minority workers' productivity, i.e., they may place more weight on negative than on positive experiences when updating their beliefs. Finally, I derive predictions regarding belief updating and hiring of workers from different minority groups, and regarding spillovers of hiring experiences across minority groups.

2.1 Model

Consider a group of workers who have observable group identity $g \in \{W, B, H\}$. Group W represents the majority, while groups B and H represent two distinct disadvantaged minorities. Adapting from Holmström (1999), the individual output from any worker that belongs to group g is given by $y_g = x_g + \varepsilon$, where x_g and ε represent measures of unobserved worker productivity and random production noise, respectively. Worker productivity is drawn from a normal distribution $\mathcal{N}(\mu_g, \sigma_g^2)$, with mean productivity $\mu_g \geq 0$ and variance σ_g^2, while the random production noise is normally distributed $\mathcal{N}(0, \sigma_\varepsilon^2)$ with mean 0 and variance σ_ε^2.[10]

[9] Recent studies that do not find evidence of labor market discrimination include Darolia et al. (2016); Decker et al. (2015); Holzer and Ihlanfeldt (1998); Kleykamp (2009); and Lahey and Mosquera (2024).

[10] I assume that the random production noise is the same across the different groups of workers.

Both workers and employers believe that x_g is uncorrelated with ε.

Workers from different groups are equally productive on average, $\mu_W = \mu_B = \mu_H$, although this is not known to employers. Having the same average productivity across workers from different groups guarantees that, on average, employers will have the same hiring experience when hiring from multiple groups. In reality, and depending on the task, there could be differences in performance across groups, causing employers to have lower initial beliefs about minority workers' performance.[11] These differences could lead employers to have lower initial beliefs about minority workers' performance. However, guaranteeing the same average hiring experience across the different groups shuts down underlying differences in productivity across the worker groups as a potential mechanism for differences in hiring,[12] allowing me to focus on employers' perceptions about each group and how these perceptions evolve over time with hiring experiences.

2.1.1 Employers' Beliefs

In a given time period, employers decide whom to hire between a pair of workers: one randomly chosen majority worker (W) and one randomly chosen minority worker (either B or H).[13] Employers make two sets of hiring decisions. In each set, the minority group is fixed to them. Without loss of generality, I focus on a case in which employers choose between W and B workers.

Before starting her hiring, at $t-1$, employer i holds beliefs about the average productivity of workers from each group, $\mathbb{E}_{i,t-1}(x_g)$, $\forall g \in \{W, B, H\}$.[14] During set t, employer i makes hiring decisions over multiple rounds. In each one, she first chooses from which group she wants to hire a worker, she is then matched with a randomly selected worker from that group, and observes the worker's output. Employers' payoffs depend on the hired worker's output. The employer-worker match is dissolved at the end of the round, and a new hiring decision follows. Hence, employers observe a different output realization for each hiring period. I consider each set of hiring decisions to be made of 10 hiring rounds. The number of output realizations observed in each set from each group of workers (in this case, W and B) depends on the distribution of hiring decisions. If an employer hires k workers from group W and $10-k$ from group B, she

[11] I employ a math task. Black and Hispanic students routinely have lower scores on the math section of the SAT than White students. See the SAT Suite of Assessments Annual Report 2024.

[12] Relaxing the assumption of equal productivity would not change the qualitative predictions of the model, since hiring decisions follow employers' expectations. However, having differences in average worker productivity would yield differences in hiring even for employers who have accurate beliefs.

[13] Even though, in reality, employers would likely choose between multiple groups at the same time, it is informative and realistic to look at pair-wise choices. Some situations that are captured by pair-wise choices include "resident" and "incoming" minorities, in which one of the minorities is available initially, and employers hire from them before a second group arrives and becomes available for hire.

[14] Since employers form beliefs about average group productivity, x_g, and $y_g = x_g + \varepsilon$, then $\mathbb{E}_t(y_g) = \mathbb{E}_t(x_g)$. Hence, priors about average group productivity match priors about average group output.

would observe k output realizations from W and $10-k$ output realizations from B. Without loss of generality, I assume that employers hire at least one worker from each group to generate exposure and therefore belief-updating.[15]

Employers update their beliefs about the average productivity of each group of workers g based on the observed signals, i.e., the average realization of outcomes. Let $y_{iB,t}$ be the average realization of B workers' output observed by employer i during the first set of hiring decisions. Conditional on observing $y_{iB,t}$, employer i forms the following posterior beliefs about group g's productivity:

$$\underbrace{\mathbb{E}_{i,t+1}\left[x_g|y_{iB,t}\right]}_{i\text{'s posteriors about } g} = \underbrace{\mathbb{E}_{i,t}[x_g]}_{i\text{'s priors about } g} + \underbrace{\frac{Cov_i(x_g, y_B)}{Var_i(y_B)}}_{\substack{\text{marginal impact of} \\ B \text{ output surprise}}} \underbrace{\left(y_{iB,t} - \mathbb{E}_{i,t}\left[x_B\right]\right)}_{B \text{ output surprise}} \tag{1}$$

In equation (1), the mean-deviation term $(y_{iB,t} - \mathbb{E}_{i,t}\left[x_B\right])$ measures employer i's output surprise from B workers, i.e., the difference between the observed realized output and the prior about the group's average productivity. Without a surprise, i.e., when the employer observes the same average output as what she expected from B workers, posterior beliefs remain the same as the initial priors, $\mathbb{E}_{i,t}[x_g]$. With a positive or a negative surprise, i.e., whenever the employer observes higher or lower average output than expected from B workers, beliefs about the productivity of group g are revised upward or downwards to the extent that the employer can associate such surprise with group g's productivity by the coefficient $\frac{Cov_i(x_g,y_B)}{Var_i(y_B)}$.[16] This coefficient represents employer i's marginal response (or the relative precision perceived) by considering the relationship between the signals from one group's output, y_B, and another group's productivity, x_g, adjusted for the informativeness of the signal. The less precise the signal is perceived, or, the less related to each other the groups g and B appear to the employer, the lower the marginal response.

While employers are observing output by workers from different groups, they might not process all the information in the same way. Specifically, it is possible that employers underreact to positive experiences with minority workers, systematically responding more strongly to negative surprises than to positive surprises, leading to a more pessimistic view of minority workers. This pessimism is consistent with confirmation bias, where employers who ex-ante believe that minority workers are less productive place more weight on information that reinforces such perception. To allow employers to have asymmetric belief updating, I adapt equation (1)

[15] If, for instance, employers only hire W workers, they would not observe any output realizations from B workers. Not hiring limits the information that employers can use to update their beliefs about B and, therefore, limits spillovers on other groups. In my experiment, only 9% of employers never hire from one of the groups.

[16] Note that this coefficient is equivalent to the slope $\hat{\beta}$ if the employer ran the following linear regression: $y_{g,t+1} = y_{g,t} + \beta(y_{B,t} - \mathbb{E}_t[x_{B,t}]) + \varepsilon$, where $E[\varepsilon|y_B] = 0$. I get back to this parallel with a regression framework in Section 4.1.

to incorporate reference-dependent preferences (Kőszegi and Rabin, 2006):

$$(2) \qquad \mathbb{E}_{i,t+1}\left[x_g|y_{iB,t}\right] = \begin{cases} \mathbb{E}_{i,t}[x_g] + \lambda_{ig}(y_B) \cdot (y_{iB,t} - \mathbb{E}_{i,t}\left[x_B\right]) & \text{if} \quad y_{iB,t} \geq \mathbb{E}_{i,t}\left[x_B\right] \\ \mathbb{E}_{i,t}[x_g] + \eta_{ig} \cdot \lambda_{ig}(y_B) \cdot (y_{iB,t} - \mathbb{E}_{i,t}\left[x_B\right]) & \text{if} \quad y_{iB,t} < \mathbb{E}_{i,t}\left[x_B\right] \end{cases}$$

where $\lambda_{ig}(y_B) = Cov_i(x_g, y_B)/Var_i(y_B)$. Under this specification, $\mathbb{E}_{i,t}\left[x_B\right]$ acts as a reference point. When the employer experiences a positive surprise ($y_{iB,t} \geq \mathbb{E}_{i,t}\left[x_B\right]$), her posteriors about the productivity of group g are updated positively with weight λ_{ig}. When the employer experiences a negative surprise ($y_{iB,t} < \mathbb{E}_{i,t}\left[x_B\right]$), her posteriors are updated negatively with weight $\eta_{ig}\lambda_{ig}$. The parameter $\eta_{ig} \geq 1$ is similar to a loss-aversion parameter, showing how negative experiences carry a higher weight than positive experiences in belief updating about group g.

2.1.2 Baseline: Belief Updating for the Same Group

When an employer updates her beliefs about group B's average productivity following hiring experiences with B workers, equation (2) can be re-written as:

(3)

$$\mathbb{E}_{i,t+1}\left[x_B|y_{iB,t}\right] = \begin{cases} (1 - \lambda_{iB}(y_B)) \cdot \mathbb{E}_{i,t}[x_B] + \lambda_{iB}(y_B) \cdot y_{iB,t} & \text{if} \quad y_{iB,t} \geq \mathbb{E}_{i,t}\left[x_B\right] \\ (1 - \eta_{iB} \cdot \lambda_{iB}(y_B)) \cdot \mathbb{E}_{i,t}[x_B] + \eta_{iB} \cdot \lambda_{iB}(y_B) \cdot y_{iB,t} & \text{if} \quad y_{iB,t} < \mathbb{E}_{i,t}\left[x_B\right] \end{cases}$$

Under this alternative specification, the employer is represented as a Bayesian learner who forms posteriors $\mathbb{E}_{i,t+1}\left[x_B|y_{iB,t}\right]$ with weights $\lambda_{iB}(y_B) = \sigma^2_{iB}/\left(\sigma^2_{iB} + \sigma^2_{\varepsilon}\right)$.[17] That is, the Bayesian employer's posterior about group B is a convex combination of their prior and the average hiring experience, with weights proportional to the perceived precision of each component. In particular, the weight placed on hiring experiences decreases with the perceived variance of B workers' output and with the random production noise variance. This means that, if an employer perceives the output realizations by B workers to be less precise, she would consider the signal to be less informative, thus placing more weight on priors and less weight on hiring experiences.

Several attributes could affect the perceived precision or informativeness of hiring experiences, in particular, the strength of the initial priors. For instance, an employer who has strong initial priors due to extensive previous experiences with members from a particular minority group might not consider recent hiring experiences as informative as an employer who has not had that much prior interaction with that minority group. Formally, if employer i has had previous interactions with or a high degree of exposure to B workers as opposed to employer

[17] Appendix A provides the detailed analytical solution

j, then employer i might consider hiring experiences less informative ($\sigma^2_{jB} > \sigma^2_{iB}$). Consequently, employer i will place lower weight on hiring experiences while forming their posterior beliefs. Another possibility is that strong initial priors are due to employers holding deep-rooted (mis)perceptions of certain groups, for instance, taste-based discrimination or prejudice, leading them to attribute much of the hiring experiences that do not conform to their priors to random noise rather than the actual productivity of the workers. Hence, they place greater weight on their priors, leading to higher values of σ^2_ε.

2.1.3 Spillovers: Belief Updating for a Different Group

Hiring experiences with workers from group B also affect beliefs about productivity of workers from other groups. In particular, when an employer updates her beliefs about group H's average performance, equation (2) can be re-written as:

$$\mathbb{E}_{i,t+1}\left[x_H | y_{iB,t}\right] = \begin{cases} \mathbb{E}_{i,t}[x_H] + \lambda_{iH}(y_B)\left(y_{iB,t} - \mathbb{E}_{i,t}\left[x_B\right]\right) & \text{if} \quad y_{iB,t} \geq \mathbb{E}_{i,t}\left[x_B\right] \\ \mathbb{E}_{i,t}[x_H] + \eta_{iB} \cdot \lambda_{iH}(y_B)\left(y_{iB,t} - \mathbb{E}_{i,t}\left[x_B\right]\right) & \text{if} \quad y_{iB,t} < \mathbb{E}_{i,t}\left[x_B\right] \end{cases} \tag{4}$$

where $\lambda_{iH}(y_B) = Cov_i(x_H, y_B)/\left(\sigma^2_{iB} + \sigma^2_\varepsilon\right)$. Even though these equations cannot be written under the alternative specification representing a Bayesian learner as in the baseline, the parameter λ_{iH} still holds the Bayesian intuition. Higher values represent higher weights placed on the hiring surprise. This weight decreases with the perceived variance of B workers' output and the random production noise variance, while it increases with the perceived covariance between B workers' output and H workers' productivity.

In addition to the perceived precision or informativeness of hiring experiences, belief updating across groups can also be affected by the perceived relationship between the groups' productivity. Consider employer i, who sees groups B and H as highly related, and employer j, who thinks they are completely different groups. Then, employer j does not consider hiring experiences with group B to be very informative about H, and will place less weight than employer i on the hiring experiences with B workers, i.e., $Cov_i(x_H, y_B) > Cov_j(x_H, y_B)$. Alternatively, an employer j may have very strong perceptions (i.e., priors) of group H workers' productivity. Experiences with workers from a different group might not change such perceptions, also leading to less weight being placed on hiring experiences with B workers, as compared to an employer with weaker priors about group H.

2.1.4 Employers' Hiring Choice

I consider risk-neutral employers whose expected utility depends on the expected profits generated by the hired workers. I adapt the setting first introduced by Lepage (2024), where

employers are not only uncertain about the individual productivity of a worker before hiring but also about the average productivity of the worker's group. Employers learn about the productivity of minority groups through hiring, update their beliefs and make hiring decisions based on previous experiences. Hiring decisions are determined by the employer's perceptions of the different groups of workers available for hire. Consider an employer i who, although unfamiliar with minority workers due to social or historical factors, has very strong initial priors, translating into lower priors about B and H workers' productivity than about W workers' productivity. Thus, employer i would be more willing to hire W workers by expecting higher output realizations.[18]

Since employers make a series of hiring decisions instead of a one-shot decision, I consider that employers hire according to a standard probabilistic choice function in which they are more likely to hire workers from the group they expect to provide higher profits. In the case in which employers choose between B and W workers, the choice rule can be written as:

$$\delta^B_{i,t+1} = \frac{\mathbb{E}_t\left[\pi_i(x_B)\right]}{\mathbb{E}_t\left[\pi_i(x_B)\right] + \mathbb{E}_t\left[\pi_i(x_W)\right]} \tag{5}$$

Where $0 \leq \delta^B_{i,t+1} \leq 1$ and $\delta^B_{i,t+1} + \delta^W_{i,t+1} = 1$. If an employer believes that workers from both groups provide the same profits, then $\delta^B_{i,t+1} = \frac{1}{2} = \delta^W_{i,t+1}$. This is, she would hire B and W workers in the same proportion. Alternatively, if she expects higher profits from W workers than from B workers, then she will hire more W workers than B workers: $\delta^B_{i,t+1} < \frac{1}{2} < \delta^W_{i,t+1}$.[19]

2.1.5 Impact of Hiring Experiences

After hiring workers from group B during set t, employers update their beliefs about the productivity of B workers following equation (3) if they experience "hiring surprises." A positive (negative) hiring surprise occurs if the observed output from the hired workers from group B, $y_{B,t}$, is above (below) the employer's expected productivity of the group before hiring, $\mathbb{E}_t[x_B]$. These hiring surprises determine positive and negative hiring experiences: an employer has a positive hiring experience if hiring leads to a positive surprise, and a negative hiring experience if hiring leads to a negative surprise.

At the end of set t, employers have updated beliefs about each group's expected productivity, $\mathbb{E}_{i,t+1}[x_B]$ and $\mathbb{E}_{i,t+1}[x_W]$, which yield updated beliefs about the expected profits and thus a new choice rule for the second set of hiring decisions, $\delta^B_{i,t+2}$. This implies that the choice rule

[18] One extension of the model is allowing employers to be risk averse, with utility function $U_i(y_g) = -exp(-r_i \cdot y_g)$, where r_i is the employer's risk-aversion degree. Assuming that the worker's output is normally distributed, with employer's beliefs about the mean $\mathbb{E}_{i,t}(y_g)$, the expected profits of hiring g workers can be written as $\mathbb{E}_t\left[U_i(y_g)\right] = -exp\left\{-r_i\left(\mathbb{E}_{i,t}(x_g) - \frac{1}{2}r_i \cdot Var_i(y_g)\right)\right\}$.

[19] Risk aversion in employers would change the choice rule to be written in terms of expected utility, following the function depicted in Note 18, rather than expected profits.

for the second set of hiring decisions could be different from $\delta^B_{i,t+1}$, the choice rule for the first set, depending on whether the employer's beliefs about workers from group B and group W are updated upward or downward, which, in turn, depends on both hiring experiences and priors' strength, as discussed in Section 2.1.2.

Following equation (4), employers update their beliefs about the productivity of workers from group H through their hiring surprises with workers from group B. In the case in which workers from group H are available for hire instead of workers from group B in the second set, the employers' choice rule would determine the proportion of H workers to be hired in the second set, $\delta^H_{i,t+2}$. As discussed in Section 2.1.3, the hiring surprise with workers from group B, alongside the strength of priors about workers from group H and the perceived covariance between the different worker groups' productivity, determine changes in posterior beliefs about the productivity of workers from group H, and thus the choice rule $\delta^H_{i,t+2}$.

2.2 Predictions

The set-up of the model makes it clear that positive and negative hiring experiences – defined relative to priors about a group's productivity – matter for belief updating and subsequent hiring decisions. I derive testable predictions for the effects of hiring experiences on posterior beliefs and subsequent hiring of workers from the same and other groups. Predictions 1 to 3 are about beliefs and belief updating, while Predictions 4 to 6 are about hiring.

Prediction 1 (Bias in beliefs against minorities):
Employers exhibit bias in beliefs against minority workers.

Although Prediction 1 does not come directly from the model, it follows the existing empirical evidence on biases against Black and Hispanic workers. Prediction 1 states that, before any hiring, employers believe that the performance of workers from group W was above the performance of workers from groups B and H: $\mathbb{E}_0(x_W) > \mathbb{E}_0(x_B)$ and $\mathbb{E}_0(x_W) > \mathbb{E}_0(x_H)$.[20]

Prediction 2 (Previous experiences matter and confirmation bias):
2a) Positive (negative) hiring experiences lead to an increase (decrease) in beliefs about the productivity of workers from the same group.
2b) Employers are pessimistic about minority workers' performance.

On the one hand, Prediction 2a follows directly equation (3) when $\lambda_g(y_g) = \frac{\sigma_g^2}{\sigma_g^2+\sigma_\varepsilon^2} > 0$ for all g. This condition holds as long as $\sigma_g^2 > 0$ and σ_ε^2 does not converge to infinity. Intuitively, if employers consider hiring experiences to be informative and not due to random noise, having

[20] This prediction also follows in-group bias perspective, as W employers might expect W workers to work harder (Dickinson et al., 2018; Banuri et al., 2022).

a positive hiring experience leads to higher posteriors about the group's productivity. The opposite occurs for negative experiences since observing productivity lower than the prior leads to lower posteriors about the group's productivity. On the other hand, Prediction 2b implies $\eta_g \geq 1$ for $g \in \{B, H\}$ in equations (3) and (4). Negative experiences with workers from a minority group lead to persistent negative biases by decreasing hiring and, therefore, learning. This leads to employers placing more weight on negative than on positive hiring experiences when updating their beliefs about minority workers' productivity.[21] Moreover, Prediction 2b is related to confirmation bias, as employers expect minority workers to perform badly. Hence, employers pay special attention to information confirming their preexisting beliefs, underreacting to positive hiring experiences (Nickerson, 1998; Hjort et al., 2021).

Prediction 3 (Spillovers of experiences in beliefs):
3a) Experiences with workers from one group may affect beliefs about the productivity of workers from another group.
3b) Spillover effects in beliefs may be asymmetric.

Prediction 3a depends on the sign of $Cov_i(x_k, y_g)$ in equation (4). When $Cov_i(x_k, y_g) \neq 0$ for $k \neq g$ and $g \in \{B, H\}$, employers update their beliefs about the productivity of workers from group k following hiring experiences with workers from group g. There are positive spillovers when $Cov_i(x_k, y_g) > 0$ and negative spillovers when $Cov_i(x_k, y_g) < 0$, since positive (negative) hiring experiences with workers from group g lead to an increase (decrease) in beliefs about the productivity of workers from group k. If instead $Cov_i(x_k, y_g) = 0$, as discussed in Section 2.1.3, there are no spillovers: employers consider the two groups to be independent or unrelated to each other. Prediction 3b implies that $Cov_i(x_k, y_g)$ may differ from $Cov_i(x_g, y_k)$, leading to asymmetric spillovers: hiring experiences with workers from group g may have a different impact on beliefs about the productivity of workers from group k than the impact of experiences with workers from group k on beliefs about the productivity of workers from group g.

Prediction 4 (Hiring gap against minorities):
Minority workers are hired less often than majority workers.

Prediction 4 follows Prediction 1 and equation (5), the hiring choice rule. Since I expect employers to have lower priors about minority workers' productivity than about majority workers' productivity, it leads to a hiring gap against minority workers. Since $\mathbb{E}_0(x_W) > \mathbb{E}_0(x_B)$ and $\mathbb{E}_0(x_W) > \mathbb{E}_0(x_H)$, then $\delta_1^W > \frac{1}{2} > \delta_1^B$ and $\delta_1^W > \frac{1}{2} > \delta_1^H$. In other words, employers hire more W workers than B workers and H workers.

[21] Prediction 2b shares intuition with Lepage (2024) on greater impacts of negative experiences, although the mechanism follows greater weight on the belief updating process instead of less hiring and learning.

Prediction 5 (Experiences matter in hiring):
Positive (negative) hiring experiences increase (decrease) hiring of workers from that group.

Prediction 5 follows Prediction 2a and equation (5), the hiring choice rule. In addition to $\lambda_g(y_g) = \frac{\sigma_g^2}{\sigma_g^2+\sigma_\varepsilon^2} > 0$ for all g, it is possible to establish from the hiring choice rule that $\frac{\partial \delta_{i,t+1}^g}{\partial y_g} > 0$, since $\frac{\partial \mathbb{E}_t[\pi_i(x_g)]}{\partial y_g} > 0$. Intuitively, if employers consider hiring experiences to be informative, positive hiring surprises will lead to higher posteriors about the group's expected productivity, leading to higher expected profits, and thus to an increase in the hiring of workers from the same group. The opposite occurs for negative experiences since hiring and observing worker performance lower than the expectation leads to lower posteriors, lower expected profits, and thus a decrease in hiring of workers from the same group.

Prediction 6 (Spillovers of experiences in hiring)
6a) Experiences with workers from one group may affect the hiring of workers from another group.
6b) Spillover effects in hiring may be asymmetric.

Prediction 6a follows Prediction 3a. Hiring experiences with workers from group g impact the hiring of workers from group k once they become available to hire through the change in beliefs about their productivity. If hiring experiences with workers from group g impact beliefs about the productivity of workers from group k positively, there would also be positive spillovers in hiring. If instead there are no spillovers in beliefs ($Cov_i(x_k, y_g) = 0$), there would be no spillovers in hiring. Similarly, Prediction 6b follows from Prediction 3b, implying that spillovers in hiring might be asymmetric.

3 Experiment Design

I test the model's prediction using an online experiment that generates a labor market where White employers can hire workers from either the majority group (White workers) or groups of Black or Hispanic workers. I conduct my experiment in two stages, as it is common in online hiring experiments (Bohren et al., 2025; Coffman et al., 2021; Eyting, 2022; Gupta, 2025; Lepage, 2024). Stage 1 collects performance data on a first group of participants, the 'workers.' Stage 2 consists of the main experiment, in which a second group of participants, the 'employers,' make hiring decisions involving the first group of participants.

The use of an online hiring experiment is ideal for a number of reasons. First, the experimental methodology allows me to vary the pool of workers (in particular, the workers available to hire) to cleanly identify how experiences with workers from one disadvantaged minority group affect beliefs about performance and hiring decisions about workers from a different disadvan-

taged minority group. Second, through the experiment, I am able to measure employers' prior beliefs about the performance of White, Black, and Hispanic workers, and examine how such priors get updated following hiring experiences. Third, by employing an online experiment, as opposed to a laboratory experiment, I am able to recruit a diverse set of subjects in the role of workers (i.e., White, Black, and Hispanic), and a geographically representative set of White employers. Fourth, online experiments provide a higher degree of anonymity compared to physical laboratory experiments, making participants more comfortable in answering sensitive questions, for instance, expressing biases against minority groups and making hiring decisions according to their true preferences. I provide more details on the advantages of an online sample and the platform I used in Section 3.5.

In what follows, I provide details on the experimental design (Sections 3.1 and 3.2) and the treatment manipulations (Section 3.3). I then discuss some design choices (Section 3.4). I conclude by describing the sampling and implementation procedures (Section 3.5).

3.1 Workers

I recruited White, Black, and Hispanic participants in the role of workers. The survey was advertised on the online platform Prolific as a 15-minute study on "Work and Hiring Decisions."[22] Upon consent, participants start the study with a demographic survey. They were then told they had been assigned to the role of workers and they had to perform a task that would generate additional bonus earnings based on their performance. The task consists of solving up to 10 multiple-choice math questions with unique answers in 10 minutes, earning them \$0.10 per correct answer. After seeing an example question, participants engage with the task. Once the 10 minutes have passed (or all the 10 questions are answered), they are thanked for their participation and informed of the total number of correct questions and their bonus payment.

The main purpose of this first stage is to create a pool of workers from three different groups, one majority (White) and two disadvantaged minorities (Black and Hispanic), to be available for hire in the second stage, the hiring experiment. Stage 1 generates workers' performance data and demographic information that I can reveal in Stage 2 to the participants in the role of employers. By creating this pool of workers from real participants of different racial and ethnic backgrounds, I do not employ deception at any point in my experiment. Each worker randomly matched with an employer in Stage 2 corresponds to an actual participant who solved the math task. This allows my second group of participants, the employers, to make incentivized hiring decisions over available workers.

[22] I provide more details on the platform and the implementation in Section 3.5. I included the study advertisement in Appendix C.

3.2 Employers

I recruited only White participants in the role of employers. The survey was advertised on Prolific as a 20-minute study on "Work and Hiring Decisions." Upon consent, participants are asked a series of demographic questions. They are then informed they have been assigned the role of employers and are randomly assigned to one of four treatments where they are asked to make hiring decisions involving a previous set of participants in the role of workers. Employers earn bonus payments in the form of tokens throughout the experiment. Tokens are converted to US dollars at the end of the study.[23]

Employers are told that workers were assigned to three different groups based on their demographics: one for White/Caucasian American workers, another one for Hispanic/Latino workers, and a third one for Black/African American workers. I disclose race or ethnic groups in a subtle way, i.e., I tell employers that the previous participants were assigned to three different groups: Group Purple, Group Orange, and Group Green. In particular, White workers are assigned to Group Purple, whereas Hispanic and Black workers are assigned to Groups Orange and Green. Hence, employers make their hiring decisions between a "worker from Group Purple" and a "worker from Group Orange" instead of choosing between a "White worker" and a "Hispanic worker." Even though the composition of the groups is common knowledge for employers, their options are over colored groups, limiting the extent to which employers explicitly choose between racial and ethnic groups, which could be subject to social desirability bias.[24]

Employers make two sets of 10 hiring decisions each (20 hiring decisions in total). In each decision, they choose from which group they want to hire a worker and their payment depends on the hired worker's performance. Specifically, they get paid 100 tokens per each problem solved correctly by the hired worker. Employers only have two groups of workers available at a time: the majority (White workers) and one of the minorities (either Black or Hispanic workers). After choosing the worker group, employers are informed about the performance of a randomly chosen worker from that group and their earnings from that hiring decision. A new hiring decision follows, where the employer faces the same majority and minority groups of workers. At the end of the experiment, one of the 10 hiring decisions in each set is randomly chosen for payment.

Before the hiring task, I elicit incentivized beliefs about worker performance. I tell employers that workers had 10 minutes to solve up to 10 multiple-choice math questions and the average number of problems solved correctly by the entire workforce, combining the three worker groups. Employers are reminded about the worker groups' composition and then asked to give their

[23] I provide more details on average participant earnings in Section 5.1.
[24] I discuss social desirability bias in Section 5.4.

guesses of the average number of problems solved correctly by members from each group. I incentivize beliefs by telling employers that they will be asked to make their guesses again later in the study, and one of their guesses will be randomly chosen and they get a bonus payment if they are correct. They get additional earnings for each correct guess. I elicit beliefs again in the middle of the hiring task (between the first and second sets) and at the end of the hiring task (after the second set). Figure 1 summarizes the experiment sequence.

Following the model, this design allows me to measure employers' initial priors, as well as the posteriors after each hiring set. The experiment also allows me to identify positive and negative hiring surprises. Therefore, I am able to test the predictions regarding initial beliefs and belief updating (Predictions 1 to 3), as well as hiring decisions (Predictions 4 to 6).

At the end of the hiring experiment, participants engage in two incentivized tasks to measure individual-level ambiguity aversion and risk aversion. Subjects complete the ambiguity aversion task introduced by Gneezy et al. (2015), in which they make two choices parallel to the one-urn Ellsberg paradox (Ellsberg, 1961). Finally, subjects complete the Eckel and Grossman (2002)'s risk aversion task, in which they make one choice over six lotteries. Appendix C contains the instructions for both tasks.

3.3 Treatments

I implement four different conditions, where I experimentally manipulate the minority group available to hire for employers. In two treatments, which I call the *Baseline* treatments, employers are presented with the same minority group throughout the two sets of hiring decisions. In the other two treatments, the *Spillover* treatments, employers are presented with a different minority group in each set of hiring decisions. Figure 1 presents the timeline of the experiment by treatment.

3.3.1 Baseline Treatments: HH and BB

In the Baseline treatments, employers can hire either a majority worker (White) or a worker from a given minority group in both sets of hiring decisions. Specifically, employers assigned to the HH treatment choose between White and Hispanic workers in both sets of hiring decisions. On the contrary, employers assigned to the BB treatment choose between White and Black workers in both sets. Employers in both HH and BB treatments make hiring decisions between Group Purple and Group Orange. The change is which specific racial or ethnic group is associated with the Orange color. Employers in the HH treatment are informed that Hispanic/Latinos belong to Group Orange, while employers in the BB treatment are informed that Black/African American workers belong to Group Orange. In both cases, employers make 20 hiring decisions involving workers from Group Purple and workers from Group Orange and learn the performance of a

randomly chosen worker from the selected group before the next hiring choice. These treatments serve as controls for the Spillover treatments by allowing me to measure beliefs and hiring patterns when only workers from the majority and one minority are available for hire.

Even though, in these treatments, employers can only hire from one minority group, they know that there are three groups of workers - Purple, Orange, and Green. This allows me to elicit employers' initial beliefs (i.e., priors) about the performance of all three groups of workers, no matter which groups employers have the chance to hire from. In fact, in all treatments, initial beliefs are elicited after the employers are informed about the hiring task and the three groups of workers, but before they are told which groups they will be able to hire from. This design choice allows me to study how employers update their beliefs about certain groups, even if they are not available for hire, as stated in the belief updating equations, (3) and (4).

3.3.2 Spillover Treatments: BH and HB

In the Spillover treatments, different minority groups are available for hire in the first and the second sets of hiring decisions. Specifically, employers choose whether to hire from Group Purple or Group Orange in the first set, and between Group Purple and Group Green in the second set. As in the Baseline treatments, the colors of the groups available for hire remain constant, but the minority groups associated with groups Orange and Green vary by treatment.

In the BH treatment, employers make hiring decisions between White and Black workers in Set 1, and between White and Hispanic workers in Set 2. This treatment, together with baseline treatment HH, allows me to study spillovers of hiring experiences from Black to Hispanic workers in the second set of the experiment. Similarly, in the HB treatment, employers make hiring decisions between White and Hispanic workers in Set 1, and between White and Black workers in Set 2. Thus, together with the baseline treatment BB, the HB treatment allows me to test how hiring of Black workers in the second set is affected by hiring experiences with Hispanic workers in the first set.[25]

3.4 Discussion of Design Choices

The hiring experiment most closely simulates a work environment where there are workers from multiple groups, but workers from the majority group are expected to perform better than minority workers. However, as there are no differences in the average performance across worker groups, I focus on the role of personal hiring experiences in the evolution of beliefs and hiring. By design, employers generate additional earnings based on the performance of the hired

[25]Employers in both BH and HB treatments make hiring decisions between Group Purple and Group Orange in Set 1, and between Group Purple and Group Green in Set 2. Employers in the BH treatment are informed that Hispanic/Latinos belong to Group Green, while employers in the HB treatment are informed that Black/African American workers belong to Group Green.

worker. Therefore, employers should choose workers from the group they expect to perform better. Note that employers only get feedback about the performance of the workers they hire, mirroring scenarios where employers do not learn about the performance of the workers they do not hire. Hence, the only information available to employers is their own past hiring experiences, so I expect personal experiences to play a primary role in belief updating and subsequent hiring decisions. By eliciting beliefs about worker performance for the three groups, regardless of the workers available for hire, I am able to assess the extent to which employers' hiring experiences with workers from one group impact beliefs about the performance of workers from a different group.

I limit employers' hiring choices to pair-wise comparisons between workers from the majority (and thus, preferred) group and workers from one of the disadvantaged minority groups to measure biases against each minority group individually. By comparing the *Baseline* treatments, BB and HH, I am able to compare biases against Black and Hispanic workers. The *Spillover* treatments simulate environments where workers from a "resident" minority are initially available to hire before workers from an "incoming" minority arrive and become available to hire. By comparing the *Spillover* with the *Baseline* treatments, I am able to assess the impact of past hiring experiences with workers from one minority group on biases against workers from a different minority group. In particular, by examining hiring decisions during Set 2 for the HB and BB treatments, I can study the impact of previous hiring experiences with Hispanic workers on the hiring of Black workers, while examining decisions for the BH and HH treatments allows me to examine the impact of hiring experiences with Black workers on the hiring of Hispanic workers.

I take several measures to improve data quality. First, as discussed earlier, I disclose race by assigning colors to each racial or ethnic group. While employers are informed that workers are assigned to groups Purple, Orange, and Green based on their demographics, they state beliefs about performance and make hiring decisions over workers belonging to groups of different colors, as opposed to explicitly disclosed races and ethnicity. This way, I aim to limit social desirability bias in employers' beliefs and hiring decisions.[26] Second, I include comprehension questions about the hiring task, the monetary incentives, and the group composition. These questions ensure that employers understand how their payment will be determined and which workers belong to which group. Third, before eliciting beliefs, I have employers attempt three of the 10 math problems that workers had to solve, using the same incentives that workers had: \$0.10 (100 tokens) per correct answer.[27] This serves two primary purposes. First, it simulates a real-life context in which employers know the job they are hiring workers for, and second,

[26] To control for color preferences, the minority group that employers encounter in the first set of hiring decisions is always associated with Group Orange.

[27] I purposely chose one easy, one medium, and one hard question (80%, 48%, and 23% of workers got correct, respectively).

it helps employers form beliefs about workers' average performance. Finally, I include three attention checks in the experiment. The first one is after the demographic survey and before the hiring task instructions, the second one is after the first belief elicitation and before the first set of hiring decisions, and the third one is after the second set of hiring decisions, before the risk and ambiguity measures. Participants are aware that their submissions might get returned if they fail attention checks.[28]

3.5 Sampling and Implementation

The experimental procedure was pre-registered on AsPredicted (#107386).[29] I recruit survey participants from the online platform Prolific for both stages of the experiment.[30] Prolific has been shown to produce high-quality data comparable to laboratory experiments (Gupta et al., 2021; Peer et al., 2021). Prolific strictly forbids researchers from asking participants for their identifiable information. This enforced anonymity protects participants' privacy and it is especially important for research on sensitive topics, as it minimizes experimenter demand effects and social desirability bias. Additionally, recent evidence shows that the size of the experimenter demand effect in online survey experiments is small; the knowledge of the experiment's purpose has no detectable effect on participants' behaviors (de Quidt et al., 2018; Mummolo and Peterson, 2019).

For the worker experiment, I only recruited subjects who met the following conditions: 1) currently in college or with a college degree, 2) between 18 and 40 years old, 3) currently reside in the United States, and 4) identify as White, Black, or Hispanic.[31] I target participants that represent a diverse workforce composed of one majority group with 150 workers (White) and two minority groups with 50 workers each (Black and Hispanic). In particular, I target college students or participants with a college degree to simulate high-skilled workers, which allows me to minimize variance in performance across the different racial and ethnic groups. Even though the recruitment on each ethnic and racial group is based on self-reported information of the participants across their profiles on Prolific, I double-check by asking them the group they identify with in the survey.

For the employer experiment, I only recruited subjects who met the following conditions: 1) 25 years old or older, 2) identify as White, 3) at least completed high school, and 4) Currently reside in the United States. I target participants who could represent the population making hiring decisions in real life. Hence, I only recruit participants who self-identify as White and who completed at least high school. I provide more details on the sample characteristics and

[28] See Prolific's Attention and Comprehension Check Policy.
[29] In Appendix D, I provide details on the pre-registered experiment design and deviations from it.
[30] In Section 5.1, I discuss characteristics of both worker and employer samples.
[31] I recruited three different samples, one for each ethnic or racial group.

earnings in Section 5.1.

I recruited nearly 2,000 subjects to participate in the experiment, either as workers (250) or as employers (1,600). The experiment was programmed in Qualtrics. All subjects were paid \$2 for completing the survey, and had the chance to earn up to \$3.30 of bonus payment.[32] The average completion time is 13 minutes for workers and 18 minutes for employers. The actual hourly earning translates into around \$12/hour for both workers and employers, which is considerably high for online surveys. In terms of data quality, 92.6% of the employers passed the three attention checks. Subjects who failed two or more attention checks were excluded from the study.

4 Estimation Strategy

My empirical strategy proceeds in two steps. I first analyze the data on beliefs and test how belief updating is affected by hiring experiences with Black or Hispanic workers. I then study how hiring experiences affect future hiring of workers from the same group or a different group. My outcomes of interest are: i) beliefs about minority workers' performance and ii) the hiring of minority workers (Black or Hispanic), following hiring experiences involving workers from the same and from a different minority group. In Appendix D, I provide details on the pre-registered analysis and when and why I deviated from it.

In what follows, I explain how I analyze the data to test the model's predictions discussed in Section 2.2 about beliefs and belief updating (Section 4.1) and about hiring (Section 4.2).

4.1 Analysis of Beliefs and Belief Updating

To test Prediction 1, I estimate employers' initial priors about workers' performance by estimating the following equation using Ordinary Least Square (OLS) regressions on the full sample of White employers:

$$Prior_{ig} = \beta_0 + \beta_1 Black_g + \beta_2 Hispanic_g + \varepsilon_i \tag{6}$$

where $Prior_i$ is the initial belief about the group g's average performance, $Black$ is a dummy equal to 1 if the group being considered is that of Black workers and 0 if it is the group of White workers. $Hispanic$ is a dummy equal to 1 if the group being considered is the group of Hispanic workers, and 0 if it is the group of White workers. The constant term β_0 represents the mean prior about White workers' performance, while β_1 and β_2 capture differentials for

[32] The maximum bonus payment for workers was \$1.10, while the maximum payment for employers was \$3.30. While workers' payments were explained in US dollar amounts, employers' payments were explained using tokens, with an exchange rate of 1,000 tokens for \$1.00.

priors about Black and Hispanic workers' performance, respectively, when compared to White workers. $\beta_1 < 0$ and $\beta_2 < 0$ imply that priors about Black and Hispanic workers are lower than priors about White workers, supporting Prediction 1.

I then focus on the belief updating process to test Prediction 2a, following a Bayesian learning framework, as stated in Section 2.1.2, by estimating the following equation:[33]

$$Posterior_{igT} = \beta_1 Prior_{igT} + \beta_2 Experience_{igT} + \varepsilon_{iT} \tag{7}$$

where $Posterior_{igT}$ is the employer's belief about group g's average performance after a set T of 10 hiring decisions. $Experience_{igT}$ refers to the average performance of the hired workers from group g in set T. The estimated coefficients β_1 and β_2 hold Bayesian intuition – a Bayesian employer's posterior will be a convex combination of their prior and the average performance of the hired workers, weighting according to the perceived precision of each component. Hence, β_1 and β_2 represent the weights placed on their prior and hiring experience, respectively. If employers have strong priors about workers' performance, hiring experiences would not affect their beliefs, then β_1 should be 1, and β_2 should be 0, i.e., the posterior should be equal to the prior. In the opposite case, when employers value hiring experiences to their maximum degree, the posterior would be equal to the hiring experience. I restrict the analysis to employers who hire at least one worker from group g in each set of hiring decisions.[34]

I estimate equation (7) independently for Black and Hispanic workers and for each instance of belief updating (meaning after the first set of hiring decisions and after the second set), pooling the BB and BH treatments and the HH and HB treatments to examine the posteriors regarding Black and Hispanic workers, respectively, after the first set. When analyzing the posteriors about the performance of a given group of workers (Black or Hispanic) after the second set of hiring decisions, I restrict the analysis to the baseline treatments HH and BB.

To examine whether belief updating differs for Black and Hispanic workers, I also estimate an extended version of equation (7) with the full sample, including interaction terms with a Hispanic treatment dummy. I report the estimates generated by the pooled specification in Appendix B.

I then test Prediction 2b, i.e., whether employers are pessimistic about minority workers' performance. This is equivalent to studying whether employers underreact to positive experiences with minority workers and, therefore, respond more to negative experiences relative to positive experiences. I follow the empirical approach first introduced by Hjort et al. (2021) and

[33] Equation (3) of the conceptual framework describes the belief updating process disaggregating the prior belief and the hiring surprise. An equivalent regression would be $Posterior_{igT} = \gamma_1 Prior_{igT} + \gamma_2(Experience_{igT} - Prior_{igT}) + \varepsilon_{iT}$. However, estimating this equation is equivalent to estimating equation (7), where $\beta_1 = \gamma_1 - \gamma_2$ and $\beta_2 = \gamma_2$. I chose to estimate equation (7) for ease of exposition and interpretation.

[34] 93.75% (92.25%) of the employers hired at least one Black (Hispanic) worker during set 1.

estimate the following equation for Black and Hispanic workers separately:[35]

$$Posterior_{igT} = \beta_1 Prior_{igT} + \beta_2 \left(Experience_{igT} - Prior_{igT}\right) + \beta_3 \left(Experience_{igT} - Prior_{igT}\right) \times NegativeSurprise_{igT} + \varepsilon_{iT} \tag{8}$$

where $NegativeSurprise_{igT} = \mathbf{1}\{Experience_{igT} - Prior_{igT} < 0\}$ is a dummy equal to 1 if the employer experienced a negative surprise in a set of hiring decisions, i.e., if the average performance of the hired workers from group g is smaller than the employer's prior, and 0 otherwise. $\beta_3 > 0$ would imply that employers place more weight on negative experiences than on positive experiences, supporting Prediction 2b, while $\beta_3 < 0$ would imply that they place more weight on positive experiences.

Predictions 3a and 3b concern the existence of spillovers in beliefs from one minority group to another. Specifically, Prediction 3a determines that hiring experiences with workers from a minority group may affect beliefs about the performance of workers from a different minority group, while Prediction 3b determines that results might be asymmetric. To test these predictions, I estimate the following equation:

$$Posterior_{ikT} = \beta_1 Prior_{ikT} + \beta_2 (Experience_{igT} - Prior_{igT}) + \varepsilon_{iT} \tag{9}$$

where k and g now refer to different minority groups. While the overall interpretation of equation (9) is similar to the interpretation of the coefficients in equation (7), a non-zero β_2 coefficient would provide support for Prediction 3a, by indicating that hiring experiences with workers from minority group g, vis-a-vis the employer's prior about that group, affect beliefs about the performance of workers from minority group k. I conduct the analysis for Black and Hispanic workers separately, following experiences with Hispanic and Black workers (in treatments HB and BH), respectively. Different β_2 coefficients from these analyses would provide support for Prediction 3b, showing that spillovers vary by minority group.

4.2 Analysis of Hiring Decisions

To test whether employers are less likely to hire Black and Hispanic workers as opposed to White workers (Prediction 4), I estimate the following equation:

$$Hire_{it} = \beta_0 + \beta_1 Minority_{it} + \varepsilon_{it} \tag{10}$$

where $Hire_{it}$ is a dummy variable equal to 1 if the worker was hired and 0 otherwise, and

[35] Hjort et al. (2021) discuss how an alternative specification would be to estimate equation (7) interacting *Prior* and *Experience* with a dummy for negative experience. However, including the term *Experience* − *Prior* and the interaction term provides an easier interpretation.

$Minority_{it}$ is a dummy equal to 1 if the worker was a minority worker, and 0 if he or she was a White worker. I estimate equation (10) separately for Black and Hispanic minority workers.

To test Prediction 5, regarding the immediate impact of hiring experiences on additional hiring of minority workers, I follow the empirical approach introduced by Lepage (2024):

$$SubsequentHiring_{igt} = \beta_0 + \beta_1 Experience_{igt} + \beta_2 Prior_{ig} + \gamma_t + \varepsilon_{it} \tag{11}$$

where $SubsequentHiring_{igt}$ is the number of workers from group g hired after period t, i.e., in the remaining $10 - t$ periods of the hiring set. $Experience_{ig,t}$ is the performance (number of problems correctly solved) of the hired worker at period t. I include γ_t, a vector of time-fixed effects. β_1 captures the impact of the performance of the minority worker hired at time t on the hiring of minority workers in the following periods. I also assess the impact of positive and negative hiring experiences, independently conducting an alternative estimation:

$$SubsequentHiring_{igt} = \beta_0 + \beta_1 Positive_{ig,t} + \beta_2 Negative_{ig,t} + \beta_3 Prior_{ig} + \gamma_t + \varepsilon_{it} \tag{12}$$

where $Positive_{ig,t}$ and $Negative_{ig,t}$ are indicator variables for positive and negative hiring experiences, respectively. I estimate equations (11) and (12) independently for each minority group. I focus on the subsequent number of hired workers through a panel specification instead of a cross-section with the total number of minority workers hired to address endogeneity issues. While the experience that employers observe with each minority worker they hire is randomly drawn and irrespective of their hiring history, the decision to hire a minority worker and observe a productivity draw beyond the first is endogenously determined by previous experiences. Hence, I hold constant the previous number of minority workers hired across employers. In other words, I form a panel such that period $t = 1$ is the first time each employer hired a minority worker. This allows me to test how experiences affect hiring more generally with an employer's previous number of hires.

Finally, I examine spillovers in hiring across minority groups (Predictions 6a and 6b) by employing a differences-in-differences specification:

$$MinorityHiring_{it} = \beta_1 Spillover_i + \beta_2 Set2_t + \beta_3 Spillover_i \times Set2_t + \delta \mathbf{X}_i + \gamma_t + \varepsilon_{it} \tag{13}$$

where $MinorityHiring_{it}$ is a dummy variable equal to 1 if employer i hired a minority worker in period t, $Spillover_i$ is a dummy equal to 1 if the employer was assigned to one of the spillover treatments: HB for Black hiring following Hispanic hiring, and BH for Hispanic hiring following Black hiring. $Set2_t$ is a dummy equal to 1 for the second set of hiring decisions, which is the relevant hiring set for the spillover treatments. The set of controls $\mathbf{X}$ contains the employer's priors about the average performances of White, Black, and Hispanic workers. $\beta_3 > 0$ would

indicate evidence of positive spillovers across minority groups. Specifically, it would suggest that employers are more likely to hire a worker from a given minority group (e.g., Black) in the second set if they had hiring experiences with a different group of minority workers (e.g., Hispanic) in the first set, supporting Prediction 6a. I estimate equation (13) separately for Black and Hispanic workers, clustering the standard errors at the individual employer level, and including hiring period (within-set) fixed-effects. If these analyses generate different coefficients for β_3, it would imply that spillovers are asymmetric, supporting Prediction 6b.

In addition to the analysis described above, I conduct various robustness checks to address some concerns that might threaten the validity of my design and results. I first conduct a placebo experiment with no reference to race or ethnicity, and then replicate the analysis with interaction terms. I also identify participants subject to experimenter demand effect, inattention, and replicate the main analysis excluding such participants. Finally, I address social desirability bias. I provide more details on the robustness checks in Section 5.4.

5 Results

In this section, I present the results of the hiring experiment. Section 5.1 reports descriptive statistics and balance tests. Section 5.2 reports the employer results for beliefs and belief updating. Section 5.3 reports the employer results for hiring. Section 5.4 reports robustness checks, including a placebo experiment with no references to race or ethnicity (Section 5.4.1). Section 5.5 reports additional results, including heterogeneous effects based on employers' initial biases (Section 5.5.2). Finally, Section 5.6 addresses mechanisms.

5.1 Descriptive Statistics and Balance Tests

Table 1 reports the demographic characteristics of the employer sample, overall and by treatment. The last column of the table reports the p-value of an ANOVA test against the null hypothesis that employers across the four treatment conditions are not jointly different from each other. Overall, employers are around 43 years old and they are left-leaning, with a political index of 0.79 out of 1. Although I screened participants in Prolific who identify as White, I included a question in the demographic survey about their racial or ethnic identity, allowing them to choose all the categories they identify with. Only 0.5% of the sample do not identify as White, and 1% identify as White and other categories.[36] 98% of the employers were born in the US, and only 4% spoke a language other than English when growing up. The average employer is college-educated and has an annual household income above $60,000. Overall, all

[36] 9 employers do not identify as White. Three identify as Hispanic, one identifies as Black, one identifies as Native Hawaiian/Pacific Islander, two identify as 'Other', and two prefer not to say.

employers' demographics are balanced across the treatment conditions, with the exception of being born in the US, which ranges from 96 percent in the HH treatment to 99 percent in the BH treatment.

In order to assess whether my sample is representative of the American White population, I compare the summary statistics presented in Table 1 with population estimates from the US Census and the American Community Survey.[37] My sample appears to be representative with respect to age, gender, income, and employment status. However, as a result of my sampling criteria, my sample is more educated and has a higher average income. In terms of political affiliation, the sampled individuals are more likely to be Independent and less likely to be Democrat but similarly likely to be Republican than the general U.S. population.[38] Finally, Figure 2 displays the geographical location of employers in my sample. Further tests indicate that the regional distribution of my sample is representative of the US population.

The demographic characteristics of the worker sample are reported in Panel A of Appendix Table B.1. Workers, on average, are around 28 years old and college educated, with 99% of the sample holding a bachelor's degree or higher. Overall, the sample of workers is balanced across all demographic characteristics except for age and gender. On average, Hispanic workers are 2 years younger (26 years old versus 28 years old for White and Black workers). 47% of White workers are women, followed by 42% of Hispanic workers and 30% of Black workers. Panel B of Appendix Table B.1 summarizes the workers' performance in the math task. The average performance is 5 out of 10 problems across all groups, with non-statistically significant differences in the average performance across groups and its variance.[39] Having the same average productivity across groups allows me to guarantee equal average experiences to employers regardless of the group they are hiring workers from. It also allows me to minimize the role of mechanisms deviating from experiences to explain biases or spillovers across minority groups, such as actual statistical differences across groups.

5.2 Employer Results: Beliefs

Figure 3 and Panels A and B of Table 2 summarize employers' initial beliefs about workers' performance in the task, for the full sample and by treatment. Overall, employers overestimate White workers' performance and underestimate Black and Hispanic workers' average performance, underestimating Hispanic workers' performance more. Panel C of Table 2 summarizes the distribution of employers according to their initial beliefs, overall and by treatment assignment. 45% of the employers are biased against both Black and Hispanic workers, meaning they

[37] See the U.S. Census QuickFacts 2021 and the American Community Survey (ACS) 2023.
[38] See the General Social Survey (GSS) 2022.
[39] The test on the equality of standard deviations generate the following p-values: for White vs. Black, p-value of 0.3910; for White vs. Hispanic, p-value of 0.9602; for Black vs. Hispanic, p-value of 0.4669

stated lower beliefs for both Black and Hispanic workers' performance. 30% of the employers have initial correct beliefs, and only 5% of the employers are biased against White workers, meaning they stated lower beliefs for White workers' performance. Estimates from equation (6), displayed in Appendix Table B.2, show an average bias in beliefs of 1 problem against Hispanic workers and 0.95 problems against Black workers, with this difference being statistically significant ($p = 0.052$). This implies that White employers believe that Hispanic and Black workers performed 19% and 18% worse than White workers, respectively, providing evidence in support of Prediction 1. My first result follows.

Result 1. Initial Beliefs

Before any hiring, White employers believe that, on average, Hispanic and Black workers performed worse than White workers, with Hispanic workers performing the lowest.

I start the analysis of belief updating in the first and second set of hiring decisions by using binned scatter plots in Figure 4. The y-axis plots the magnitude and direction of updating ($Posterior - Prior$). The the x-axis displays the hiring surprise ($Experience - Prior$), i.e., the difference between the average performance of the hired minority workers and the prior the employer had at the beginning of the hiring set. I draw a few lessons from the left panel of Figure 4, which presents belief updating after the first set of hiring decisions. First, the relationship between belief updating and the hiring surprise appears to be linear, in line with the Bayesian framework. Second, there does not seem to be any difference in the slope for positive and negative surprises, suggesting symmetric updating. Third, the slopes obtained for Black and Hispanic workers are not different, suggesting no differences in the weights placed by employers on priors and experiences when forming their posteriors on each group of workers.

The right panel of Figure 4 plots belief updating in relation to hiring surprises in the second set of hiring decisions. Compared to the first instance of belief updating, the slope tends to be flatter for both groups, suggesting that the weight placed on priors increases over time. In addition, the figure shows a weaker updating response for Hispanic workers than for Black workers. In fact, there seems to be no response (flat line) to positive surprises for Hispanic workers, while there is still negative updating following negative surprises.

The regression estimates confirm these findings. Panel A of Table 3 presents the OLS estimates for equation (7) described in Section 4.1, and Figure 5 plots the estimated coefficients. Consistent with the Bayesian framework, $\hat{\beta}_1$ and $\hat{\beta}_2$ are both positive and statistically significant, and they add up to approximately 1. Columns 1 and 3 of Table 3 show that employers place approximately equal weights (one-half weights) on their priors and hiring experiences - defined as average performance of hired minority workers - when forming posterior beliefs about Hispanic and Black workers' performance after the first set of hiring decisions. For Hispanic workers, the weight placed on priors increases in the second belief updating instance, with em-

ployers placing about two-thirds of the weight on their prior. Specifically, the weight placed on hiring experiences falls from a 52.9% weight (Set 1) to a 32.4% weight (Set 2), a 39 percent decrease. In the case of Black workers, there is also an increase in the weight placed on the prior in the second set of hiring decisions (column 4 versus column 3). However, the weight placed on hiring experiences with Black workers only falls by 26 percent from the first to the second set, falling from a 56.5% weight (Set 1) to a 41.6% weight (Set 2).

Pooling all the data and including interaction terms for the minority group being Hispanic as opposed to Black, confirms that there are no differences across minority groups in the weight placed on priors versus experiences after Set 1 (see Appendix Table B.4). However, when considering the second set of hiring decisions, the regression analysis indicates that employers place an additional 11% weight on their priors and a 10% less weight on their experience when updating their beliefs about Hispanic workers as compared to Black workers.

Overall, the scatter plots and regression estimates show that, although the weight given to hiring experiences may differ by minority group and by hiring set, employers tend to always update their beliefs positively in response to positive hiring experiences and negatively in response to negative hiring experiences, confirming Prediction 2a.

Result 2a. Belief Updating

Positive (negative) hiring experiences lead to an increase (decrease) in posterior beliefs about the performance of workers from the same group.

To test whether positive and negative surprises play a different role in belief updating and whether employers respond more to negative experiences (pessimism), I estimate equation (8). Panel B of Table 3 shows no evidence of asymmetric updating, with the exception of column 2, which refers to belief updating regarding Hispanic workers' performance in the second hiring set. The estimates suggest that when employers have hiring experiences with Hispanic workers in the second set, they place 22.7% more weight on negative experiences ($p = 0.078$) than on positive experiences. Specifically, employers place a 23% weight on positive or neutral hiring experiences, while they place a 45.7% weight on negative experiences. This is consistent with the right panel of Figure 4, and leads to Result 2b:

Result 2b. Pessimism toward Hispanic workers

Employers are pessimistic about Hispanic workers' performance but not about Black workers' performance.

In summary, the comparison of priors and belief updating regarding Black and Hispanic workers reveals that White employers have higher priors about Black workers' performance, and they learn more about them by placing more weight on their hiring experiences when forming posteriors. White employers believe that Hispanic workers have the lowest average

performance, and they are pessimistic about their performance. These beliefs are harder to modify with hiring experiences, making the bias toward Hispanic workers stickier over time.

Next, I examine spillovers in beliefs across minority groups. Specifically, I study whether White employers update their beliefs about the performance of workers from one minority group (e.g., Black) after hiring experiences with workers from a different minority group (e.g., Hispanic). In column 1 of Table 4, I report estimates of equation (9) when focusing on Hispanic workers' performance, i.e., I test whether and to what extent employers learn from hiring surprises concerning Black workers (a different minority group) when updating their beliefs about Hispanic workers. Column 2 replicates the analysis for Black workers' performance. The estimates show that White employers place a 17% weight on their hiring surprises with Black workers when forming their posteriors about Hispanic workers, and a 13% weight on hiring surprises with Hispanic workers when forming their posteriors about Black workers.

These coefficients indicate that hiring experiences with workers from one minority group affect the beliefs about the performance of workers from a different minority group, i.e., there exist spillovers in beliefs across minority groups. Further analysis, displayed in Panel B of Table 4, shows that employers place equal weights on positive and negative experiences with Black workers when forming their posteriors about Hispanic workers' performance. In contrast, they turn optimistic about Black workers following hiring experiences with Hispanic workers. Specifically, White employers place around a 20% weight on positive experiences with Hispanic workers when forming their posterior beliefs about Black workers, but they do not place any weight on negative experiences with Hispanic workers. In other words, employers seem to only update their beliefs about Black workers if they had positive experiences with Hispanic workers, confirming Predictions 3a and 3b, i.e., spillovers in beliefs are asymmetric. Result 3 follows.

Result 3. Spillovers in Beliefs

Employers update their beliefs about Hispanic workers' performance after positive and negative hiring experiences with Black workers, but only update their beliefs about Black workers' performance after positive hiring experiences with Hispanic workers.

5.3 Employer Results: Hiring

I start the analysis of hiring decisions by examining whether Hispanic and Black workers are less likely to be hired than White workers. The coefficient estimates of equation (10), displayed in Appendix Table B.3, show that Hispanic and Black workers are 17.6 and 14.1 percentage points less likely to be hired than White workers, respectively,[40] confirming Prediction 4.

Result 4: Hiring Differentials

[40]The difference between these two coefficients is not statistically significant, as shown in Column 3.

Hispanic and Black workers are less likely to be hired than White workers.

I then test the immediate impact of positive and negative hiring experiences on subsequent hiring of workers from a given group during the first set of hiring decisions. Table 5 presents estimates of equations (11) and (12), where the dependent variable is the number of hired minority workers following a minority hire. There is a clear positive relationship between the performance of the hired minority worker and the number of minority workers hired over the following periods. In the case of Hispanic workers, employers hire 0.061 additional workers per each additional problem solved by the hired worker ($p = 0.000$). In other words, if the hired Hispanic worker solved 10 problems correctly, employers hire, on average, 0.61 more Hispanic workers in the following periods, which is a 26 percent increase given the control mean of 2.269 workers. In addition, Column 2 of Table 5 indicates that a positive surprise with a Hispanic worker (i.e., a worker with performance above the employer's prior) increases future hiring of Hispanic workers by 0.108 workers, while a negative surprise decreases hiring by 0.170 workers. The effects observed for Black workers are very similar (columns 3 and 4), except for positive surprises: the hiring of Black workers increases by a larger margin following a positive surprise. Employers hire 0.147 more Black workers following a positive surprise, and 0.160 fewer workers following a negative surprise. The difference between these two coefficients is not statistically significant.[41] Result 5 follows:

Result 5. Hiring Dynamics

Positive (Negative) hiring experiences with minority workers increase (decrease) hiring in the following periods.

To test whether there exist spillovers in hiring across minority groups (Prediction 6a), I turn my attention to hiring decisions in the second set, conditional on treatment assignment. In the baseline, the hiring gap against Hispanic workers decreases from 14 pp in the first set to 10.85 pp in the second set. In the case of Black workers, it actually increases from 7.6 pp to 8.1 pp. None of these changes is statistically significant, showing that the hiring of the same minority workers is relatively stable over time.

The estimated coefficients of equation (13), displayed in Table 6, show that the bias against minority workers in the second set of hiring decisions decreases only for Black workers in the HB treatment, i.e., when the hiring of Black workers follows the hiring of Hispanic workers. Specifically, Black workers are 5.1 pp more likely to be hired in the second set in the HB treatment, where employers could only hire White or Hispanic workers in Set 1, as compared to the BB treatment, where employers could hire Black or White workers in Set 1. This represents

[41] Appendix Table B.5 shows that the impacts of positive and negative surprises on future hiring are not statistically different across Black and Hispanic workers.

an 11% increase over the control mean ($p = 0.000$). In contrast, Hispanic workers are not more likely to be hired in the spillover treatment (BH treatment) as compared to the HH treatment. These results not only show how the hiring of workers from one minority group is affected by previous hiring experiences with workers from a different minority group, but also that spillover effects differ across minority groups, confirming Predictions 6a and 6b. This leads to Result 6:

Result 6. Spillovers in Hiring
There are positive hiring spillovers from Hispanic to Black workers only; Hispanic hiring is not affected by previous Black hiring.

In summary, hiring experiences with Hispanic workers make employers more likely to hire Black workers, up to the point that White and Black workers are equally likely to be hired. The asymmetric hiring spillovers can be explained by the asymmetric spillovers in beliefs discussed in Section 5.2 and a larger bias against Hispanic workers, which makes employers turn to Black workers once they become available for hire. Specifically, employers consider Hispanic workers to be less productive than Black workers. Negative hiring experiences with Hispanic workers reinforce the bias against them, and employers do not consider them to be informative about Black workers. Hence, their posterior beliefs about Black workers' performance do not change. On the contrary, positive hiring experiences with Hispanic workers are considered informative, increasing employers' posterior beliefs about Black workers' performance and thus making them more likely to hire Black workers once they become available for hire.

5.4 Robustness Checks

In this subsection, I conduct numerous robustness checks. I start by addressing the possibility that my experiment does not capture true biases related to workers' race or ethnicity. In Section 5.4.1, I report the results of a placebo experiment where employers are presented with the same colored groups of workers (Purple, Orange, and Green) but are not informed about the relationship between the colors and the race or ethnicity of the workers. In Sections 5.4.2, 5.4.3, and 5.4.4, I address concerns regarding experimenter demand effect, participants' possible inattention, and social desirability bias.

5.4.1 Placebo Experiment: Neutral Groups

Even though the asymmetric results obtained for Black and Hispanic workers suggest that White employers in the experiment based their decisions on their real perceptions of Black and Hispanic workers, one concern is that the racial and ethnic identity of workers is not made salient enough by the colored group labels. It is also possible that subjects in the role of employers did not believe that the workers belonged to real racial and ethnic minorities. In

addition, the observed employer behavior may be driven by differences in the groups' sizes (the Purple group being three times as large as the Orange and the Green groups) rather than by truthful attitudes towards Black and Hispanic workers, as opposed to White workers.

I conducted a placebo experiment on Prolific in the Summer of 2024, which employed what I call the "neutral" groups. I recruited a second sample of 400 White participants with the same demographics as the employers in the main experiment.[42] The structure and procedures of the placebo experiment are the same, except for the information given to employers about the composition of the worker groups. Employers were only told that previous participants in the role of workers were assigned to three different groups, such that 150 participants were assigned to Group Purple, 50 participants were assigned to Group Orange, and 50 participants were assigned to Group Green. No references to racial or ethnic identity or demographics of any source were mentioned when explaining how the workers were assigned into the three groups.

I conducted two treatments of the placebo experiment: a baseline and a spillover treatment. In the "neutral baseline" treatment, employers choose to hire between workers from Group Purple and Group Orange in both sets of hiring decisions, and Group Orange consists of, in fact, Hispanic workers. I therefore replicate the HH treatment with neutral groups. In the "neutral spillover" treatment, employers choose to hire between workers from Group Purple and Group Orange in the first set and between workers from Group Purple and Group Green in the second set, and Group Green consists of Black workers. I therefore replicate the HB treatment with neutral groups.

There are three main findings from the Neutral experiment. First, when employers are unaware that Orange and Green workers are in fact Hispanic and Black workers, they underestimate both minority workers' performance less, as shown in Appendix Table B.6. Second, employers in the placebo experiment update their beliefs in a more balanced way, i.e., they do not display sticky priors regarding the performance of either minority group. Appendix Table B.7 shows that employers place less weight on their priors and more weight on their hiring experiences when forming their posteriors about both minority and majority workers' performance. Third, the placebo experiment shows that without racial and ethnic identities associated with group colors, there is no evidence of spillovers in beliefs across different minority groups. Appendix Table B.8 shows that hiring experiences with workers from Group Orange do not affect beliefs about Group Green workers' performance.

In sum, the placebo experiment validates my design and confirms that the results I obtained in my hiring experiment are driven by White employers perceptions and attitudes towards White, Black and Hispanic workers.

[42]See Section 3.5 for the recruitment specifications.

5.4.2 Experimenter Demand Effect

Another concern is that my results are driven by experimenter demand effects. This would imply that subjects might have correctly guessed the true purpose of the experiment and thus acted in the way they believed the experimenter expected them to act. There are several features of my design that help address this issue. First, my experiment was conducted online. The procedure and the anonymity enforced by Prolific, such that the researcher does not have access to the participant's identifying information, should help alleviate experimenter demand effects concerns. Moreover, recent research has proven that online experiments are resistant to demand effects (de Quidt et al., 2018; Mummolo and Peterson, 2019). Second, I asked participants an open-ended question at the end of the experiment to elicit their guesses about the purpose of the study ("What do you think this study was about?"). If the participant's response contains words related to any of the following roots: discrimination, racism, bias, stereotypes, or prejudice, I identify that participant as having correctly guessed the purpose of the study. Overall, 12% of the subjects correctly guessed the purpose of my study, as shown in Appendix Table B.9, with overall balance across treatments.

As a robustness checks, I examine whether my results still hold if I exclude the 12% of subjects who correctly guessed the purpose of my study from regressions. The estimates, reported in Appendix Tables B.10, B.11, B.12, and B.13, are very similar in magnitude and statistical significance to the main estimates in Sections 5.2 and 5.3. This indicates that my findings are unlikely to be driven by experimenter demand effects.

5.4.3 Inattention

One concern with online experiments is that subjects might not pay enough attention due to lack of monitoring. As discussed in Section 3.4, I included comprehension questions and three attention checks to make sure that subjects paid attention to critical aspects of the design. Overall, 99.4% of the participants passed at least two of the three attention checks. Out of them, 93.2% passed all three attention checks. Appendix Table B.9 shows that, from the approved submissions (failed less than two attention checks), the proportion of subjects failing one attention check is largely the same across treatments. To investigate whether the 6.81% of subjects who failed one attention check biased my results, I conduct robustness checks by excluding these subjects from the analysis. The estimates reported in Appendix Tables B.10, B.11, B.12, and B.13, are very similar in magnitude and statistical significance compared to the main estimates in Sections 5.2 and 5.3. Hence, my findings are unlikely to be driven by subjects' inattention.

5.4.4 Social Desirability Bias

An additional concern is that social desirability bias affects my results, given that the study measures behaviors that could be considered socially inappropriate. For instance, it might be socially inappropriate to state that White workers' performance is higher than (or to hire White workers over) Black or Hispanic workers. Thus, social desirability bias should lower the likelihood of finding evidence of bias against Black and Hispanic workers in my experiment, making my findings conservative.

To assess the extent to which social desirability bias affected my findings, I included the Marlowe-Crowne social desirability scale in the Prolific survey, which measures a subject's propensity to act in a socially desirable way (Crowne and Marlowe, 1960). I use the 7-item X1 scale as described in Fisher (1993) and construct the index following Anderson (2008)'s method. I follow Dhar et al. (2022), testing for heterogeneous effects based on the social desirability score.[43] I report the results in Appendix Tables B.14 and B.15. The coefficients on the interaction terms are not statistically different from zero. This pattern is reassuring and shows that the stated beliefs and hiring patterns are similar for subjects with low versus high propensity to give socially desirable responses. Hence, my results are unlikely to be driven by social desirability bias.

5.5 Additional Analysis

In this Section, I present additional results. I first describe results on belief updating and hiring experiences with White workers. I then present heterogeneous effects by initial bias.

5.5.1 Results for White workers

To investigate belief updating about White workers' performance, I replicate the analysis in Section 5.2 and present the estimated coefficients in Appendix Table B.16. First, in a similar way to belief updating about minority workers' performance, employers place more weight on priors than on hiring experiences, and this tendency grows over time, suggesting sticky beliefs about White workers. On average, the weight placed on priors rises from a 53% weight during Set 1 to a 71% weight during Set 2, representing a 34 percent increase. Second, employers generally tend to place more weight on negative experiences only during the first set, although the additional weight placed on negative experiences is just a 44 percent increase compared to the weight placed on positive or neutral experiences (the additional weight placed on negative experiences with Hispanic workers is 82 percent). During the second set of hiring decisions, employers do not place any differential weight on negative experiences.

[43] Section C.3.1 presents the seven questions used to measure social desirability bias.

With respect to hiring decisions, I replicate the analysis from Section 5.3 and present the estimated coefficients in Appendix Table B.17. The results show how the performance of each hired White worker impacts the number of White workers hired in the next periods, although to a lesser extent than the impact found in minority workers. However, this does not translate into less hiring since the control mean is higher. Specifically, positive surprises increase hiring by 0.086 workers, while negative surprises decrease hiring by 0.114 workers.

5.5.2 Heterogeneous Effects by Initial Bias

As discussed in Section 5.2, 45% of the employers had lower initial priors about both Black and Hispanic workers' performance than about White workers. Since they have lower expectations, which makes them less likely to hire minority workers, they might update their beliefs in a different way than employers who are more likely to hire minority workers, and their hiring might respond more to positive and negative hiring experiences. To test for possible differences in the effects of hiring experiences on belief updating and subsequent hiring, I replicate the analysis for the two subsamples of "biased" and "unbiased" employers.

Figure B.2 presents the binned scatter plots for belief updating independently for each subsample. As a first impression, the slopes for biased and unbiased employers do not seem to be different in each belief updating instance, suggesting that biased employers do not place differential weights on experiences when forming their posterior beliefs. Moreover, the flatter slopes in updating after the second set are noticeable for both subsamples, suggesting that all the employers tend to place more weight on their priors over time. However, the flatter slope to the right of zero for the second belief updating instance about Hispanic workers' performance is noticeable for the biased subsample but not for the unbiased subsample, suggesting the pessimism about Hispanic workers is driven by biased employers. Appendix Table B.18 presents the estimated coefficients, confirming these impressions. While there are no differences in the weights placed on priors and experiences when forming their posteriors, the asymmetric updating about Hispanic workers' performance is only present for biased employers. Nevertheless, the regression estimates show that unbiased employers are pessimistic about Black workers' performance during Set 1.

Appendix Table B.19 presents the analysis for spillovers in beliefs, confirming that there are spillovers in beliefs across minority groups for both biased and unbiased employers. Panel B presents the estimates for asymmetric updating. While non-statistically significant, the estimates show similar patterns from the main analysis, with negative coefficients for the interaction term between the hiring surprise and dummy for negative surprise, and similar magnitude with the un-interacted term, suggesting only positive updating about Black workers' performance following hiring experiences with Hispanic workers for both biased and unbiased employers.

Appendix Table B.20 presents the analysis for dynamics in hiring decisions, showing how hiring experiences have greater impacts on biased employers, especially for Hispanic workers: while a positive hiring surprise with a Hispanic worker does not affect the hiring of additional Hispanic workers for unbiased employers (column 6), there is a positive and significant effect for biased employers (column 2).

Finally, Appendix Table B.21 presents the coefficients for spillovers in hiring. First, as expected, biased employers are less likely to hire minority workers, as shown by the lower control means. Second, there is a slight increase in the hiring of Hispanic workers in the baseline treatment for biased employers (column 1), but not for unbiased employers (column 3). The hiring of Black workers remains unchanged for both types of employers (columns 2 and 4). Third, there are no differences in the hiring of Hispanic workers following hiring experiences with Black workers, confirming a lack of spillovers (columns 1 and 3). On the contrary, there are positive spillovers only from Hispanic to Black workers for both subsamples (columns 2 and 4). In the case of unbiased employers, the hiring gap against Black workers turns positive: there is a preference to hire Black workers over White workers after having hiring experiences with Hispanic workers.

The heterogeneity analysis shows that, while biased employers still update their beliefs, they do not do it differently than unbiased employers, suggesting that their (biased) beliefs are not stickier. However, as expected, unbiased employers are not pessimistic about Hispanic workers' performance. Finally, the asymmetric spillovers hold for both biased and unbiased employers: Black workers are more likely to be hired by employers who could only hire White or Hispanic workers in the first set. This confirms that employers overall have a greater degree of bias against Hispanic workers, and even employers who are initially indifferent with respect to Hispanic workers, remain indifferent after hiring experiences with Black workers, while a preference in favor of Black workers emerges after hiring experiences with Hispanic workers.

5.6 Exploratory Analysis of Mechanisms

In this Section, I present an exploratory (not pre-registered) analysis of possible mechanisms behind my primary results.[44] I start by decomposing the sources of discrimination in an attempt to disentangle taste-based and statistical discrimination towards Black and Hispanic workers. I then assess whether natural exposure to Black and Hispanic individuals – which varies by the location of the study participants – drives my results.

[44] I did not pre-register the analysis of mechanisms. Appendix D presents details on how and where I do not follow the pre-registered analysis.

5.6.1 Sources of Discrimination

I aim to identify the sources of discrimination in the hiring of Black and Hispanic workers by measuring the role of (inaccurate) statistical versus taste-based discrimination. To do so, I follow Bohren et al. (2025) and compare the gap in beliefs about the performance of Black and Hispanic workers as compared to White workers, with the corresponding hiring gap.[45] In the case in which the hiring gap and the gap in beliefs about performance against minority workers are the same, the entire gap would be of an inaccurate statistical nature, without prejudice or preference for either group. If the hiring gap exceeds the gap in beliefs, the remainder of the gap would have a taste-based nature, coming from prejudice against minorities. On the contrary, when the hiring gap is less than the gap in beliefs, the remainder is also of a taste-based nature, although it indicates preference for minority workers (or against White workers).

Appendix Table B.22 presents the gaps in beliefs and hiring by treatment. For simplicity, and given that both beliefs and hiring decisions range from 0 to 10, I will refer to both gaps as measured in "points." Looking at Hispanic workers in the baseline treatment, the difference in hiring is -1.183 points, with a difference in beliefs of -1.03 points during Set 1. Thus, the remaining -0.153 point difference between the belief and hiring gap suggests prejudice against Hispanic workers. This prejudice increases to -0.342 points in Set 2. Looking at Black workers, while the decomposition also suggests prejudice against them, the magnitudes are around one-third of the prejudice against Hispanic workers: the prejudice against Black workers is -0.057 points in the first set and -0.11 during the second set. Hence, the decomposition in the baseline treatments suggests prejudice as the source of discrimination against both Black and Hispanic workers, with more prejudice against Hispanic workers. Looking at the Spillover treatments in Set 2, the gaps in beliefs in both cases are greater than the gaps in hiring, suggesting the presence of taste-based attributes in favor of minority workers. In the case of Hispanic workers (treatment BH), the hiring gap is -0.41 points, with a belief gap of -0.99 points, while in the case of Black workers (treatment HB), the hiring gap is +0.015 points, and the belief gap is -0.798. Hence, there is a 0.58 point preference in favor of Hispanic workers and a 0.813 point preference in favor of Black workers.

There are three valuable lessons from the decomposition analysis. First, the results confirm the greater bias against Hispanic workers. Second, while additional hiring experiences might decrease the gap in beliefs about minority workers' performance, stable hiring gaps lead to an increase in prejudice against minority workers over time, suggesting that employers might disguise their animus as statistical at first, but it is uncovered as prejudice over time. Third, the spillover treatments change the source of discrimination by suggesting preference in favor of

[45] Bohren et al. (2025) discusses how the traditional outcomes-based test assumes accurate beliefs by comparing the performance gap and the hiring gap. By using employers' actual beliefs, it is possible to identify the source of discrimination.

minority workers. Employers develop preferences in favor of Black workers when they become available to hire, explaining a positive hiring gap while still holding a negative difference in beliefs. Employers also develop preferences in favor of Hispanic workers in the BH treatment, although it does not compensate for the greater prejudice against them exhibited by the end of Set 1. Thus, by decomposing the sources of discrimination and identifying prejudice against (or preferences in favor of) minority workers, I find suggestive evidence that employers develop preferences in favor of minority workers following hiring experiences with workers from the other minority group. This suggests that the lack of spillovers in hiring from Black to Hispanic workers follows a larger preference in favor of Black workers and a larger prejudice against Hispanic workers.

5.6.2 The Role of Natural Exposure to Black and Hispanic individuals

As discussed in the conceptual framework (Section 2), employers' beliefs about minority workers' performance and the weights placed on priors and experiences may be based on personal experiences or exposure to minority workers. To investigate whether this is the case, I match my experiment data with U.S. Census Data on the proportion of Hispanic and Black population at the county level, provided by the National Historical Geographic Information System.[46] On average, employers live in counties where 13.4% of the population is Black and 17.5% of the population is Hispanic, which compare quite well with the national averages of 13.6% and 19.5%, respectively.[47]

First, I analyze whether the proportion of Black or Hispanic populations affects both employers' initial beliefs about workers' performance and the probability of hiring minority workers. Appendix Table B.23 shows that initial beliefs about workers' performance are not affected by the proportion of Black and Hispanic populations in the county of residence of the employers (columns 1 to 3). While natural exposure does not impact the hiring of Hispanic workers (column 4), it does impact the hiring of Black workers (column 5). Specifically, employers who live in areas with a greater percentage of Hispanic residents are more likely to hire Black workers.

As a second step, I employ threshold regressions to investigate whether and how the proportions of Blacks and Hispanics in the population affect the weights placed by employers on priors and experiences when updating their beliefs about the average performance of each group of workers. This methodology allows coefficients in a linear regression to differ across regions, given a threshold value identified by the model.[48] While belief updating about Black workers' performance is unaffected by the proportion of Blacks and Hispanics in the population, belief

[46] See the National Historical Geographic Information System (NHGIS) 2022.
[47] See U.S. Census QuickFacts for Black and Hispanic populations.
[48] I use the Stata command *threshold*, which finds the threshold values (if any) and presents the estimation coefficients in each region, defined as above or below each threshold value

updating about Hispanic workers after the second set of hiring changes with a threshold value of 0.2932, meaning that the coefficient estimates vary for employers who live in areas where the percentage of Hispanics in the population is above or below 29.32%.[49] Appendix Table B.24 shows the estimated coefficients. Employers who live in counties with a low Hispanic population according to the threshold (corresponding to 30% of the employers) place more weight on priors during the second set, placing a third-quarters weight on priors, compared to the equal weights placed by employers who live in areas with a higher presence of Hispanic individuals in the population. This suggests that White employers' beliefs about Hispanic workers' performance are stickier when they live in areas with low Hispanic presence.

The exploratory analysis provides suggestive evidence consistent with the main results. In particular, Result 6 points out that employers are more likely to hire Black workers following hiring experiences with Hispanic workers, but Hispanic hiring is not affected by previous Black hiring. This asymmetry is consistent with the existence of a higher degree of bias against Hispanic workers than against Black workers, as shown by lower expectations of Hispanic workers' performance (Result 1) and pessimism about Hispanic workers' performance, but not about Black workers' performance (Result 2b).

My exploratory analysis is also consistent with the evidence presented by Fouka and Tabellini (2022), showing that Mexican immigration improves White Americans' racial attitudes, increases support for pro-Black government policies, and lowers anti-Black hate crimes.[50] Their findings are explained through the re-categorization of Black individuals as in-group members since the inflows of Mexican immigrants, a more distant group, increase the salience of immigration status instead of race. By exploiting the variation in the percentage of Black and Hispanic individuals in the employers' county of residence, I find that natural exposure to Hispanic individuals increases the likelihood that Black workers are hired, while Hispanic hiring is unaffected by natural exposure to Black individuals. If employers see Hispanics as less productive than Black workers, experiences with them (either through hiring experiences or natural exposure) may induce employers to update their beliefs about Black workers upward - they learn that Hispanic workers are better than they thought, inducing them to have better perceptions about Black workers and thus being more likely to hire them once they become available for hire.

[49] The estimations show that the optimal number of thresholds is zero for both instances of belief updating about Black workers' performance and for belief updating about Hispanic workers' performance after the first set of hiring decisions.

[50] Fouka and Tabellini (2022) follow a self-categorization framework, hypothesizing that individuals categorize others as in-group or out-group members based on shared attributes, and the inflow of relatively more distant groups increases the salience of attributes that differ from the majority.

6 Conclusion

In this paper, I employed an online experiment to examine beliefs about performance and hiring of Black and Hispanic workers. The experiment also allowed me to test whether and how hiring experiences with workers from one disadvantaged minority (e.g., Black workers) have positive or negative spillovers on beliefs about performance and hiring of workers from another disadvantaged minority (e.g., Hispanic workers), and whether the answer depends on which minority workers are available to hire first.

My first finding is that employers exhibit biases in beliefs and hiring against both Hispanic and Black workers: they believe that White workers perform better and hire White workers more frequently than Black and Hispanic workers. My second finding is that the bias against Hispanic workers is more pronounced than the bias against Black workers. Employers have stickier beliefs about Hispanic workers' performance (placing more weight on priors than on hiring experiences over time) and are pessimistic about their performance: they place more weight on negative experiences when updating their beliefs than on positive experiences. My third finding is that there are asymmetric spillovers in beliefs and hiring. Specifically, employers update their beliefs about Hispanic workers' performance after positive and negative experiences with Black workers, but only update their beliefs about Black workers' performance after positive experiences with Hispanic workers. In other words, employers do not consider negative experiences with Hispanic workers informative when updating their beliefs about Black workers' performance. In addition, I find evidence of spillovers in hiring from Hispanic to Black workers, but not from Black to Hispanic workers. This means that while Black workers are more likely to be hired following hiring experiences with Hispanic workers, the hiring of Hispanics is unaffected by previous hiring experiences with Black workers. Finally, I provide suggestive evidence that natural exposure to Hispanics makes employers less likely to have sticky beliefs about Hispanic workers' performance and more likely to hire Black workers.

Overall, this paper shows that experiences with workers from one group could have consequences on outcomes for a different group. Importantly, while hiring experiences with Hispanic workers positively affect outcomes for Black workers, hiring experiences with Black workers do not negatively affect outcomes for Hispanic workers. This paper has implications for future work. To start with, my findings suggest that policies that promote the representation of disadvantaged minorities should depend on the targeted group, given that biases against Black and Hispanic workers are different in nature. In particular, my findings add evidence on how Hispanics are perceived as more distant than Blacks, in line with previous research showing that increased exposure to members from a distant group (either through hiring experiences or through individuals residing in the same area) could reshape intergroup relations by re-categorizing groups according to the attributes in common (Fouka and Tabellini, 2022). Hence,

hiring experiences with Hispanic workers make more salient the attributes associated with them (such as immigration status or differences in language),[51] leading to a re-categorization of Black workers as in-group members.

Future work should also address some caveats of my design. First, my paper focuses on decisions concerning Hispanic and Black workers. Subsequent research should directly examine the differential biases and stereotypes surrounding different minority groups and thus better inform policies aimed at fomenting diversity, considering that such policies could have impacts on members from other groups. Second, my paper studies decisions made from the perspective of White employers. A diverse labor market is characterized not only by a diverse set of workers but also by employers and hiring managers. Future work could test whether Black or Hispanic employers also hold biased beliefs against workers from their own groups and, hence, prefer to hire White workers. This could be done by expanding the setting of my experiment to include Black and Hispanic employers. Third, my experiment focuses on decisions made by a single employer. Future studies could investigate whether decisions made by a committee affect hiring decisions among a diverse pool of candidates and whether the racial and ethnic composition of hiring committees, in particular the inclusion of "diverse" members (i.e., Black or Hispanic), leads to (positive or negative) changes in evaluations and hiring decisions concerning all under-represented minority candidates, as opposed to just those who "match" with the profiles included in the committee.

[51] This includes general perceptions about Hispanic or Latino workers and stereotypes surrounding them that are not present with respect to Black workers. See, for instance, this article in The Economist, 2023.

References

Aaronson, D. and B. J. Phelan (2022). The evolution of technological substitution in low-wage labor markets. *Review of Economics and Statistics*, 1–45.

Aigner, D. J. and G. G. Cain (1977). Statistical theories of discrimination in labor markets. *Ilr Review 30*(2), 175–187.

Alesina, A. and E. La Ferrara (2014). A test of racial bias in capital sentencing. *American Economic Review 104*(11), 3397–3433.

Altonji, J. G. and C. R. Pierret (2001). Employer learning and statistical discrimination. *The Quarterly Journal of Economics 116*(1), 313–350.

Anderson, M. L. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects. *Journal of the American Statistical Association 103*(484), 1481–1495.

Arrow, K. J., O. Ashenfelter, and A. Rees (1973). Discrimination in labor markets. *The Theory of Discrimination*, 3–33.

Banuri, S., C. Eckel, and R. K. Wilson (2022). Does cronyism pay? Costly ingroup favoritism in the lab. *Economic Inquiry 60*(3), 1092–1110.

Bartoš, V., M. Bauer, J. Chytilová, and F. Matějka (2016). Attention discrimination: Theory and field experiments with monitoring information acquisition. *American Economic Review 106*(6), 1437–1475.

Becker, G. S. (1957). *The Economics of Discrimination*. University of Chicago press.

Bendick Jr, M., C. W. Jackson, V. A. Reinoso, and L. E. Hodges (1991). Discrimination against Latino job applicants: A controlled experiment. *Human Resource Management 30*(4), 469–484.

Benson, A. and L. P. Lepage (2024). Learning to Discriminate on the Job. *Available at SSRN 4155065*.

Bertrand, M. and E. Duflo (2017). Field experiments on discrimination. *Handbook of Economic Field Experiments 1*, 309–393.

Bertrand, M. and S. Mullainathan (2004). Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *American Economic Review 94*(4), 991–1013.

Bohren, J. A., K. Haggag, A. Imas, and D. G. Pope (2025). Inaccurate statistical discrimination: An identification problem. *Review of Economics and Statistics 107*(3), 605–620.

Burgess, S. and E. Greaves (2013). Test scores, subjective assessment, and stereotyping of ethnic minorities. *Journal of Labor Economics 31*(3), 535–576.

Campos-Mercade, P. and F. Mengel (2024). Non-Bayesian statistical discrimination. *Management Science 70*(4), 2549–2567.

Chan, A. (2022). Discrimination and Quality Signals: A Field Experiment with Healthcare Shoppers. *Unpublished manuscript*.

Coffman, K. B., C. L. Exley, and M. Niederle (2021). The role of beliefs in driving gender discrimination. *Management Science 67*(6), 3551–3569.

Crowne, D. P. and D. Marlowe (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology 24*(4), 349.

Darolia, R., C. Koedel, P. Martorell, K. Wilson, and F. Perez-Arce (2016). Race and gender effects on employer interest in job applicants: new evidence from a resume field experiment. *Applied Economics Letters 23*(12), 853–856.

de Quidt, J., J. Haushofer, and C. Roth (2018, November). Measuring and bounding experimenter demand. *American Economic Review 108*(11), 3266–3302.

Decker, S. H., N. Ortiz, C. Spohn, and E. Hedberg (2015). Criminal stigma, race, and ethnicity: The consequences of imprisonment for employment. *Journal of Criminal Justice 43*(2), 108–121.

Dhar, D., T. Jain, and S. Jayachandran (2022). Reshaping adolescents' gender attitudes: Evidence from a school-based experiment in India. *American Economic Review 112*(3), 899–927.

Dickinson, D. L., D. Masclet, and E. Peterle (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.

Eckel, C. C. and P. J. Grossman (2002). Sex differences and statistical stereotyping in attitudes toward financial risk. *Evolution and Human Behavior 23*(4), 281–295.

Ellsberg, D. (1961). Risk, ambiguity, and the Savage axioms. *The Quarterly Journal of Economics 75*(4), 643–669.

Esponda, I., R. Oprea, and S. Yuksel (2023). Seeing what is representative. *The Quarterly Journal of Economics 138*(4), 2607–2657.

Evsyukova®, Y., F. Rusche®, and W. Mill (2025). LinkedOut? A field experiment on discrimination in job network formation. *The Quarterly Journal of Economics 140*(1), 283–334.

Eyting, M. (2022). Why do we discriminate? The role of motivated reasoning.

Fisher, R. J. (1993). Social desirability bias and the validity of indirect questioning. *Journal of Consumer Research 20*(2), 303–315.

Fouka, V., S. Mazumder, and M. Tabellini (2022). From immigrants to Americans: Race and assimilation during the Great Migration. *The Review of Economic Studies 89*(2), 811–842.

Fouka, V. and M. Tabellini (2022). Changing In-Group Boundaries: The Effect of Immigration on Race Relations in the United States. *American Political Science Review 116*(3), 968–984.

Gneezy, U., A. Imas, and J. List (2015). Estimating individual ambiguity aversion: A simple approach. Technical report, National Bureau of Economic Research.

Gomez-Vasquez, D. (2025). Beyond Hiring Quotas.

Gomez-Vasquez, D., T. Salmon, and D. Serra (2025). Reaching for the Stars? An Experiment on Discrimination in Hiring.

Goncalves, F. and S. Mello (2021). A few bad apples? Racial bias in policing. *American Economic Review 111*(5), 1406–1441.

Gupta, N. (2025). Can Temporary Affirmative Action Improve Representation?

Gupta, N., L. Rigotti, and A. Wilson (2021). The experimenters' dilemma: Inferential preferences over populations. *arXiv preprint arXiv:2107.05064*.

Haaland, I. and C. Roth (2023). Beliefs about racial discrimination and support for pro-black policies. *Review of Economics and Statistics 105*(1), 40–53.

Hjort, J., D. Moreira, G. Rao, and J. F. Santini (2021). How research affects policy: Experimental evidence from 2,150 brazilian municipalities. *American Economic Review 111*(5), 1442–1480.

Holmström, B. (1999). Managerial incentive problems: A dynamic perspective. *The Review of Economic Studies 66*(1), 169–182.

Holzer, H. J. and K. R. Ihlanfeldt (1998). Customer discrimination and employment outcomes for minority workers. *The Quarterly Journal of Economics 113*(3), 835–867.

Hurst, E., Y. Rubinstein, and K. Shimizu (2024). Task-based discrimination. *American Economic Review 114*(6), 1723–1768.

Kessler, J. B., C. Low, and X. Shan (2024). Lowering the playing field: Discrimination through sequential spillover effects. *Review of Economics and Statistics*, 1–28.

Kleykamp, M. (2009). A great place to start? The effect of prior military service on hiring. *Armed Forces & Society 35*(2), 266–285.

Kline, P., E. K. Rose, and C. R. Walters (2024, August). A discrimination report card. *American Economic Review 114*(8), 2472–2525.

Kőszegi, B. and M. Rabin (2006). A model of reference-dependent preferences. *The Quarterly Journal of Economics 121*(4), 1133–1165.

Lahey, J. and R. Mosquera (2024). Age and hiring for high school graduate Hispanics in the United States. *Journal of Population Economics 37*(1), 21.

Lang, K. and A. K.-L. Spitzer (2020). Race discrimination: An economic perspective. *Journal of Economic Perspectives 34*(2), 68–89.

Lepage, L.-P. (2024). Experience-based discrimination. *American Economic Journal: Applied Economics 16*(4), 288–321.

Leung, M. D. (2018). Learning to hire? Hiring as a dynamic experiential learning process in an online market for contract labor. *Management Science 64*(12), 5651–5668.

List, J. A. (2004). The nature and extent of discrimination in the marketplace: Evidence from the field. *The Quarterly Journal of Economics 119*(1), 49–89.

List, J. A. (2006). Friend or foe? A natural experiment of the prisoner's dilemma. *The Review of Economics and Statistics 88*(3), 463–471.

McConnell, B. and I. Rasul (2021). Contagious animosity in the field: Evidence from the Federal Criminal Justice System. *Journal of Labor Economics 39*(3), 739–785.

Mummolo, J. and E. Peterson (2019). Demand effects in survey experiments: An empirical assessment. *American Political Science Review 113*(2), 517–529.

Neumark, D. (2018). Experimental research on labor market discrimination. *Journal of Economic Literature 56*(3), 799–866.

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology 2*(2), 175–220.

Ondrich, J., S. Ross, and J. Yinger (2003). Now you see it, now you don't: why do real estate agents withhold available houses from black customers? *Review of Economics and Statistics 85*(4), 854–873.

Pager, D., B. Bonikowski, and B. Western (2009). Discrimination in a low-wage labor market: A field experiment. *American Sociological Review 74*(5), 777–799.

Peer, E., D. Rothschild, A. Gordon, Z. Evernden, and E. Damer (2021). Data quality of platforms and panels for online behavioral research. *Behavior Research Methods*, 1–20.

Phelps, E. S. (1972). The statistical theory of racism and sexism. *American Economic Review 62*(4), 659–661.

Quillian, L., D. Pager, O. Hexel, and A. H. Midtbøen (2017). Meta-analysis of field experiments shows no change in racial discrimination in hiring over time. *Proceedings of the National Academy of Sciences 114*(41), 10870–10875.

Shi, Y. and M. Zhu (2023). Model Minorities in the Classroom? Positive Bias towards Asian Students and its Consequences. *The Journal of Public Economics*.

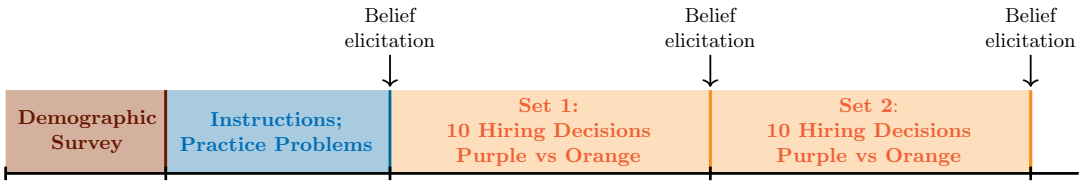
Szymanski, S. (2000). A market test for discrimination in the English professional soccer leagues. *Journal of Political Economy 108*(3), 590–603.

Wozniak, A. (2015). Discrimination and the effects of drug testing on black employment. *Review of Economics and Statistics 97*(3), 548–566.

Figures

Figure 1: Sequence of the Experiment

(a) Baseline Treatments



(b) Spillover Treatments

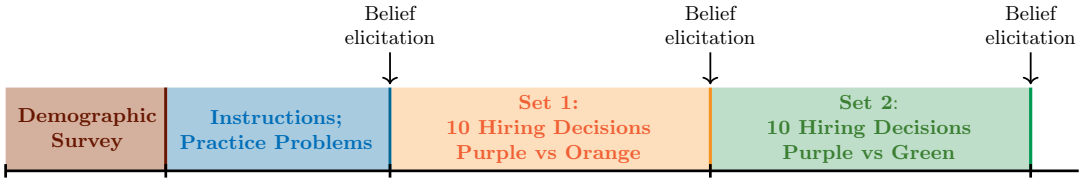
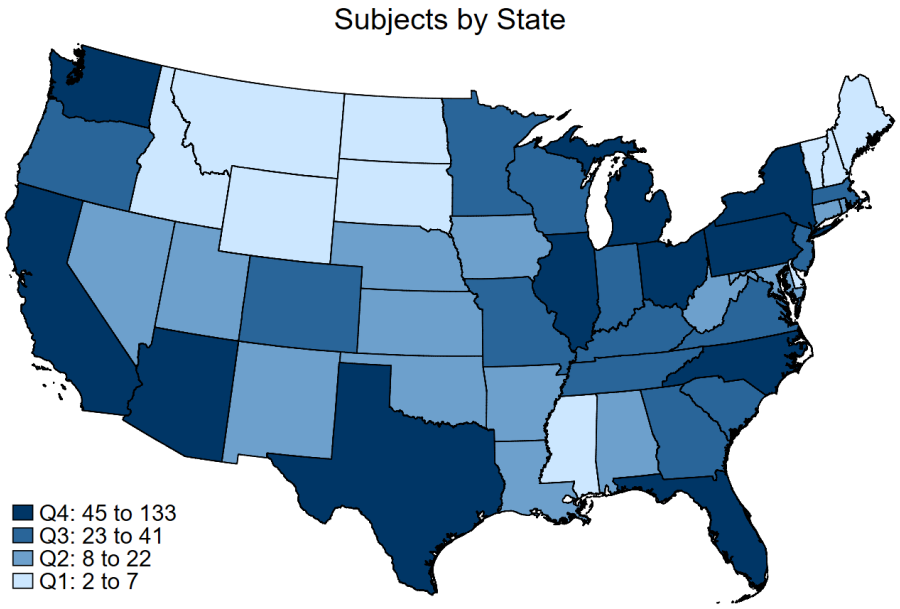
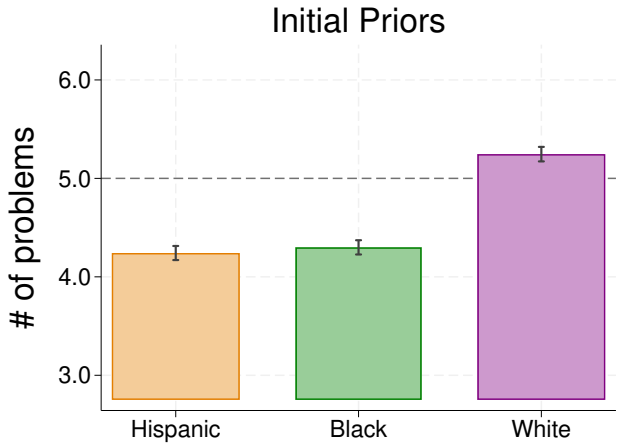


Figure 2: Geographical Location of Study Participants



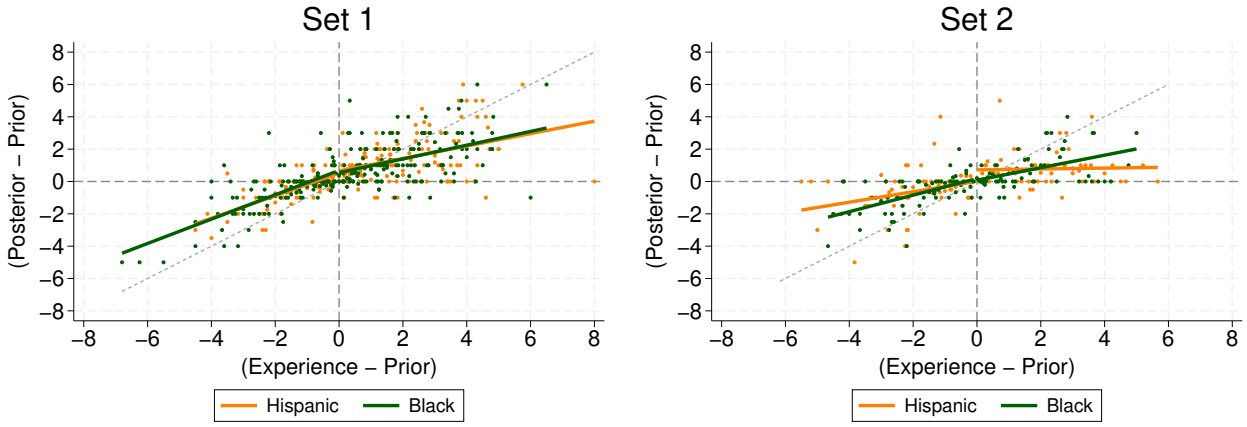
Notes: This figure shows the geographical location of the 1,600 subjects in the role of employers.

Figure 3: Initial Beliefs About Workers' Performance



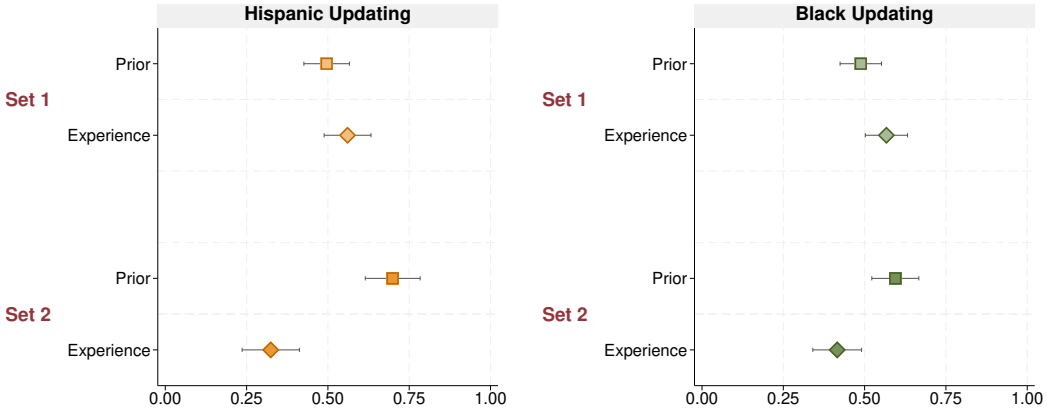
Notes: This figure shows the average initial beliefs stated by employers about each of the worker groups, pooling all the treatments (N=1,600). The first belief elicitation takes place before the first set of hiring decisions after employers are informed about the hiring task, but before they know which groups are available for hire. Confidence intervals included at the 95% level.

Figure 4: Belief Updating - Baseline Treatments



Notes: These figures compare the difference in employer's beliefs after hiring decisions (posterior beliefs minus prior beliefs) and the hiring surprise (average hiring experience minus prior beliefs), averaged over bins. *Prior* is the employer's belief about the minority group's average performance before the hiring decisions. *Experience* is the average performance of the hired workers. *Posterior* is the employer's belief about the minority group's average performance after the hiring decisions. The left panel shows statistics for the full sample in Set 1: treatments HH and HB for Hispanic workers and treatments BB and BH for Black workers. The right panel compares statistics for the second set of hiring decisions for the baseline treatments only: HH for Hispanic workers and BB for Black workers. The slope is based on a linear regression with a constant term.

Figure 5: Belief Updating



Notes: This figure presents the estimated coefficients for *Prior* and *Experience* for belief updating about minority workers for the full sample. Error bars represent 95% confidence intervals. Table 3 reports the full regression estimates.

Tables

Table 1: Employer Sample Characteristics

	All	HH	BB	HB	BH	ANOVA p-val
Age	43.44	42.77	43.74	43.78	43.46	0.681
	(13.09)	(13.07)	(13.12)	(13.04)	(13.17)	
Female	0.54	0.55	0.55	0.53	0.55	0.894
	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	
White Only	0.98	0.98	0.98	0.99	0.99	0.879
	(0.12)	(0.13)	(0.13)	(0.11)	(0.11)	
Born in the US	0.98	0.96	0.99	0.98	0.99	0.026**
	(0.14)	(0.19)	(0.11)	(0.15)	(0.10)	
Both parents born in the US	0.93	0.93	0.93	0.93	0.94	0.780
	(0.25)	(0.26)	(0.26)	(0.26)	(0.23)	
Language different than English	0.04	0.04	0.03	0.04	0.03	0.497
	(0.19)	(0.21)	(0.17)	(0.20)	(0.16)	
Student	0.08	0.07	0.09	0.09	0.06	0.517
	(0.27)	(0.26)	(0.28)	(0.28)	(0.24)	
Currently Employed	0.69	0.69	0.71	0.69	0.69	0.913
	(0.46)	(0.47)	(0.46)	(0.46)	(0.46)	
Education level	3.50	3.47	3.50	3.51	3.50	0.953
	(1.10)	(1.09)	(1.12)	(1.12)	(1.08)	
Income level	7.04	6.98	7.04	6.92	7.20	0.680
	(3.38)	(3.54)	(3.41)	(3.29)	(3.28)	
Political index	0.79	0.78	0.79	0.79	0.79	0.820
	(0.27)	(0.27)	(0.26)	(0.27)	(0.27)	
Political leaning	2.40	2.44	2.40	2.37	2.37	0.926
	(1.66)	(1.72)	(1.67)	(1.64)	(1.62)	
Social Desirability	0.00	0.00	0.01	0.01	-0.02	0.857
	(0.55)	(0.54)	(0.57)	(0.56)	(0.55)	
Willingness to take risks	4.67	4.73	4.62	4.61	4.71	0.874
	(2.51)	(2.53)	(2.58)	(2.52)	(2.41)	
Willingness to do good causes	7.02	7.03	7.08	6.87	7.10	0.678
	(2.88)	(2.81)	(2.86)	(3.03)	(2.82)	
Observations	1,600	400	400	400	400	

Notes: A total of 1,600 individuals participated in the online experiment as employers. Subjects were recruited from Prolific. This table reports the mean of each demographic variable across different treatment conditions. The corresponding standard deviation is reported in parentheses. Language different than English is an indicator variable for responding yes to "Did you speak a language different than English at home when growing up?" Education level is a categorical variable ranging from Less than high school (1) to Post-graduate degree (5). Income level is a categorical variable ranging from less than $10,000 (1) to $150,000 or more (12). Political index ranges from 0 to 1, a higher value means more left-leaning. Political leaning ranges from 0 (extreme left) to 6 (extreme right). Social desirability is an index standardized around the HH mean. Willingness to take risks is a categorical variable ranging from Completely unwilling to take risks (0) to Very willing to take risks (10). Willingness to do causes is a categorical variable ranging from Completely unwilling to do so (0) to Very willing to do so (10). The last column reports p-values from an ANOVA test against the null hypothesis that subjects across the four experimental treatment conditions are not jointly different from each other. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 2: Employer Priors

	All	HH	BB	HB	BH	ANOVA p-val
Panel A: Priors						
Priors about White	5.25	5.29	5.32	5.20	5.17	0.424
	(1.50)	(1.48)	(1.51)	(1.47)	(1.54)	
Priors about Hispanic	4.24	4.27	4.29	4.16	4.25	0.626
	(1.46)	(1.42)	(1.46)	(1.45)	(1.51)	
Priors about Black	4.30	4.22	4.40	4.29	4.29	0.358
	(1.48)	(1.55)	(1.38)	(1.53)	(1.45)	
Panel B: Initial bias						
Initial bias against Hispanic	-1.00	-1.02	-1.03	-1.04	-0.93	0.734
	(1.63)	(1.64)	(1.72)	(1.46)	(1.70)	
Initial bias against Black	-0.95	-1.08	-0.92	-0.92	-0.88	0.326
	(1.66)	(1.75)	(1.53)	(1.63)	(1.71)	
Panel C: Distribution of employers						
% Biased against Both	0.45	0.49	0.45	0.46	0.41	0.229
	(0.50)	(0.50)	(0.50)	(0.50)	(0.49)	
% More biased against Hispanic	0.26	0.24	0.29	0.27	0.25	0.366
	(0.44)	(0.43)	(0.45)	(0.44)	(0.43)	
% More biased against Black	0.22	0.26	0.22	0.20	0.20	0.172
	(0.41)	(0.44)	(0.41)	(0.40)	(0.40)	
% Correct beliefs	0.30	0.30	0.28	0.28	0.34	0.160
	(0.46)	(0.46)	(0.45)	(0.45)	(0.48)	
% Biased in favor Both	0.05	0.05	0.05	0.04	0.04	0.985
	(0.21)	(0.21)	(0.22)	(0.21)	(0.21)	
Observations	1,600	400	400	400	400	

Notes: A total of 1,600 individuals participated in the online experiment as employers. Subjects were recruited from Prolific. Panel A reports the mean of priors (initial beliefs) about each worker group's performance across different treatment variations. The corresponding standard deviation is reported in parentheses. Panel B reports the mean initial bias against minority workers, defined as the difference between the priors about minority workers' performance and priors about White workers' performance. Panel C reports the distribution of employers. *% Biased against Both* is the percentage of employers biased against both Black and Hispanic workers. *% Correct beliefs* is the percentage of employers who stated 5 as the average performance of the workers from the different groups. *% Biased in favor Both* is the percentage of employers biased in favor of both Black and Hispanic workers. The last column reports p-values from an ANOVA test against the null hypothesis that subjects across the four experimental treatment conditions are not jointly different from each other. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 3: Belief Updating

	Posteriors about Hispanic		Posteriors about Black	
	Set 1	Set 2	Set 1	Set 2
	(1)	(2)	(3)	(4)
Panel A: Belief Updating				
Prior	0.496***	0.699***	0.488***	0.594***
	(0.036)	(0.043)	(0.032)	(0.037)
Experience	0.560***	0.324***	0.567***	0.416***
	(0.037)	(0.045)	(0.033)	(0.038)
Prior Mean	4.28	4.74	4.45	4.80
Posterior Mean	4.74	4.78	4.81	4.78
Observations	727	357	731	367
Treatments	HH, HB	HH	BB, BH	BB
H_0: Prior = Experience	0.371	0.000***	0.225	0.016**
Panel B: Test for Asymmetric Updating				
Prior	1.075***	1.056***	1.067***	1.027***
	(0.014)	(0.018)	(0.014)	(0.015)
Experience - Prior	0.522***	0.230***	0.537***	0.360***
	(0.053)	(0.063)	(0.052)	(0.065)
(Experience - Prior) × Negative Surprise	0.146	0.227*	0.085	0.123
	(0.097)	(0.120)	(0.104)	(0.104)
Prior Mean	4.28	4.74	4.45	4.80
Posterior Mean	4.74	4.78	4.81	4.78
Observations	727	357	731	367
Treatments	HH, HB	HH	BB, BH	BB

Notes: This table reports OLS estimates. The dependent variables are posterior beliefs about average group performance (beliefs elicited after each set of hiring decisions). *Prior* is the employer's belief about the minority group's average performance before the hiring set. *Experience* is the average performance of the hired minority workers in each set. *Negative Surprise* is a dummy equal to 1 if the average hiring experience is less than the employer's priors about the average group performance. The estimations on Set 1 pool treatments HH and HB for posteriors about Hispanic workers' performance and treatments BB and BH for posteriors about Black workers' performance. Estimations on Set 2 use only the baseline treatments' data (HH for Hispanic workers and BB for Black workers). This analysis restricts the sample to employers who hired at least one minority worker in both sets of hiring decisions, corresponding to 727 employers in treatments HH and HB (90.8% of the sample) and 731 employers in treatments BB and BH (91.4% of the sample). Appendix Table B.4 presents the estimates pooling for Black and Hispanic workers, including a dummy variable for Hispanic workers. Robust standard errors are in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 4: Spillovers in Beliefs

	Posteriors about:	
	Hispanic	Black
	(1)	(2)
Panel A: Belief Updating		
Prior	0.961***	0.980***
	(0.011)	(0.011)
Hiring Surprise with Other Minority	0.166***	0.129***
	(0.032)	(0.037)
Prior Mean	4.32	4.31
Posterior Mean	4.29	4.41
Observations	750	738
Treatments	BB, BH	HH, HB
Panel B: Test for Asymmetric Updating		
Prior	0.951***	0.952***
	(0.014)	(0.017)
Hiring Surprise with Other Minority	0.193***	0.196***
	(0.042)	(0.053)
Hiring Surprise with Other Minority × Neg. Surp.	-0.074	-0.231**
	(0.090)	(0.116)
Prior Mean	4.32	4.31
Posterior Mean	4.29	4.41
Observations	750	738
Treatments	BB, BH	HH, HB
H_0: Surp + Surp × NegSurp = 0	0.088*	0.689

Notes: This table reports OLS estimates. The dependent variable is the posterior beliefs after the first set of hiring decisions about Hispanic workers' performance (column 1) and about Black workers' performance (column 2). *Prior* is the employer's beliefs about minority workers' performance before the first set of hiring decisions. *Hiring Surprise with Other Minority* is the difference between the average performance of the hired Black workers and the prior beliefs about Black workers' performance (column 1), and the difference between the average performance of the hired Hispanic workers and the prior beliefs about Hispanic workers' performance (column 2). *Negative Surprise* is a dummy equal to 1 if the hiring surprise is negative. The estimations on beliefs about Hispanic workers' performance pool treatments BB and BH, when employers choose to hire between White and Black workers in Set 1. The estimations on beliefs about Black workers' performance pool treatments HH and HB, when employers choose to hire between White and Hispanic workers in Set 1. The analysis restricts the sample to employers who hired at least one Black or one Hispanic worker in Set 1 (94% of employers in treatments BB and BH, and 92% of employers in treatments HH and HB). Robust standard errors are in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 5: Hiring Dynamics (Set 1)

	Hispanic Hiring		Black Hiring	
	(1)	(2)	(3)	(4)
Experience$_t$	0.061***		0.061***	
	(0.006)		(0.005)	
Positive Surprise$_t$		0.108***		0.147***
		(0.039)		(0.038)
Negative Surprise$_t$		-0.170***		-0.160***
		(0.041)		(0.039)
Prior about Minority	0.093***	0.136***	0.035	0.074**
	(0.033)	(0.033)	(0.034)	(0.034)
Prior about White	-0.063**	-0.063**	-0.067**	-0.067**
	(0.031)	(0.031)	(0.032)	(0.032)
Control Mean	2.269	2.269	2.249	2.249
Observations	3,527	3,527	3,618	3,618
Clusters (Employers)	738	738	750	750
Time FE	Yes	Yes	Yes	Yes
H_0: Pos + Neg = 0		0.408		0.850

Notes: This table reports OLS estimates of the impact of hiring experiences with minority workers on subsequent hiring of the group, i.e., the number of minority workers hired in the following periods, holding constant the number of minority workers hired across employers. *Experience* is the number of problems solved correctly by a hired minority worker. *Positive Surprise* is a dummy variable equal to 1 if the minority worker's performance is greater than the employer's priors about average performance. *Negative Surprise* is a dummy variable equal to 1 if the minority worker's performance is less than the employer's priors about average performance. Columns 1 and 2 present the estimates for hiring of Hispanic workers, pooling treatments HH and HB, when employers choose to hire between White and Hispanic workers in Set 1. Columns 3 and 4 present the estimates for hiring of Black workers, pooling treatments BB and BH, when employers choose to hire between White and Black workers in Set 1. Appendix Table B.5 presents the estimates pooling for Black and Hispanic workers, including a dummy variable for Hispanic workers. The estimations include time fixed-effects. Robust standard errors clustered at the individual level are in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 6: Spillovers in Hiring

	Pr. Hiring	
	Hispanic	Black
	(1)	(2)
Set 2	0.016	-0.002
	(0.010)	(0.010)
Spillover Tr.	0.007	-0.008
	(0.015)	(0.014)
Set 2 × Spillover Tr.	0.021	0.051***
	(0.014)	(0.015)
Priors about Hispanic	0.015**	0.020***
	(0.007)	(0.006)
Priors about Black	0.026***	0.030***
	(0.006)	(0.006)
Priors about White	-0.030***	-0.031***
	(0.005)	(0.005)
Control Mean	0.436	0.457
Observations	16,000	16,000
Clusters (Employers)	800	800
Period FE	Yes	Yes

Notes: This table reports OLS estimates using the panel of 20 hiring decisions. The dependent variable is a dummy equal to 1 if a minority worker was hired in a given time period. *Set 2* is a dummy equal to 1 for the second set of hiring decisions, i.e., periods 11 to 20. *Spillover Tr* is a dummy variable equal to 1 if the employer was assigned to the spillover treatments: HB for Black hiring and BH for Hispanic hiring. Column 1 presents the estimates for Hispanic hiring, pooling treatments HH and BH, when employers choose to hire between White and Hispanic workers in Set 2. Column 2 presents the estimates for Black hiring, pooling treatments BB and HB, when employers choose to hire between White and Hispanic workers in Set 2. The estimations include time (within set) fixed-effects. Robust standard errors clustered at the individual level are in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Appendix A Bayesian Learning

Bayes' Law. The probability of event A occurring, given that event B occurred is:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}, \quad \text{with} \quad P(B) \neq 0$$

The Bayes' Law comes from the definition of a conditional probability: $P(A|B) = P(A \cap B)/P(B)$. For continuous random variables, Bayes' law can be applied to probability densities. Let f be a continuous random variable with a smooth distribution. Then, the probability density of event A, given that event B occurred is:

$$f(A|B) = \frac{f(B|A)f(A)}{f(B)}, \quad \text{with} \quad f(B) = \int_{-\infty}^{\infty} f(B|A)f(A)dA$$

As described in Section 2, the individual output from any worker that belongs to group g is given by $y_g = x_g + \varepsilon$, where x_g measures unobserved worker productivity, $x_g \sim \mathcal{N}(\mu_{ig}, \sigma_{ig}^2)$, and ε measures random production noise, $\varepsilon \sim \mathcal{N}(0, \sigma_\varepsilon^2)$. The random production noise is conditionally independent of $\mu_{ig} - x_g$, meaning that the individual output and the expected worker performance are related only because they are informative about x_g, but their errors are independent. This independence implies $\mathbb{E}[(\mu_{ig} - x_g)(y_g - x_g)] = 0$. Given the prior belief and the observed output, the employer forms a posterior belief about x_g using Bayes' law:

$$\hat{x}_g \equiv \mathbb{E}[x_g|y_g] = \frac{\sigma_g^{-2}\mu_{ig} + \sigma_\varepsilon^{-2}y_{ig}}{\sigma_g^{-2} + \sigma_\varepsilon^{-2}}$$

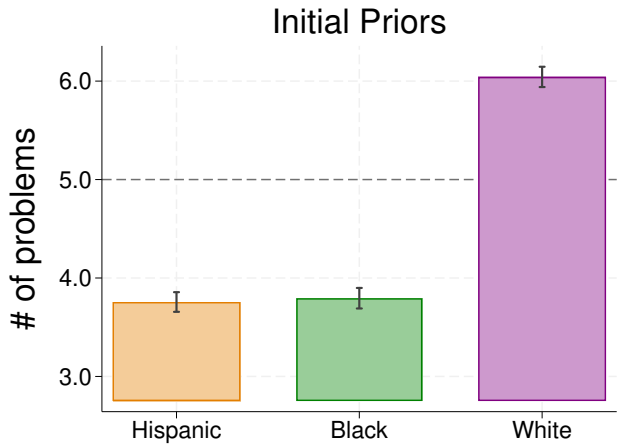
With normal random variables, the posterior belief is a weighted average of the prior belief and the observed output. Each is weighted by its relative precision. If the observed output contains no information about x_g, it would have zero precision. Hence, the posterior belief would be the same as the prior belief. When an employer hires workers from group B, with $\mu_{iB} = \mathbb{E}_t[x_B]$:

$$\begin{aligned}
\mathbb{E}_{t+1}[x_B|y_{B,t}] &= \frac{\sigma_\varepsilon^2\mathbb{E}_t[x_B] + \sigma_B^2 y_{iB,t}}{\sigma_B^2 + \sigma_\varepsilon^2} \\
&= \frac{\sigma_\varepsilon^2}{\sigma_B^2 + \sigma_\varepsilon^2}\mathbb{E}_t[x_B] + \frac{\sigma_B^2}{\sigma_B^2 + \sigma_\varepsilon^2}y_{iB,t} \\
&= \mathbb{E}_t[x_B] + \frac{\sigma_B^2}{\sigma_B^2 + \sigma_\varepsilon^2}(y_{iB,t} - \mathbb{E}_t[x_B])
\end{aligned}$$

This equation represents equation (3) from the conceptual framework, showing how employers form beliefs regarding Black workers' performance for $t+1$, given the hiring experiences during t.

Appendix B Additional figures and tables

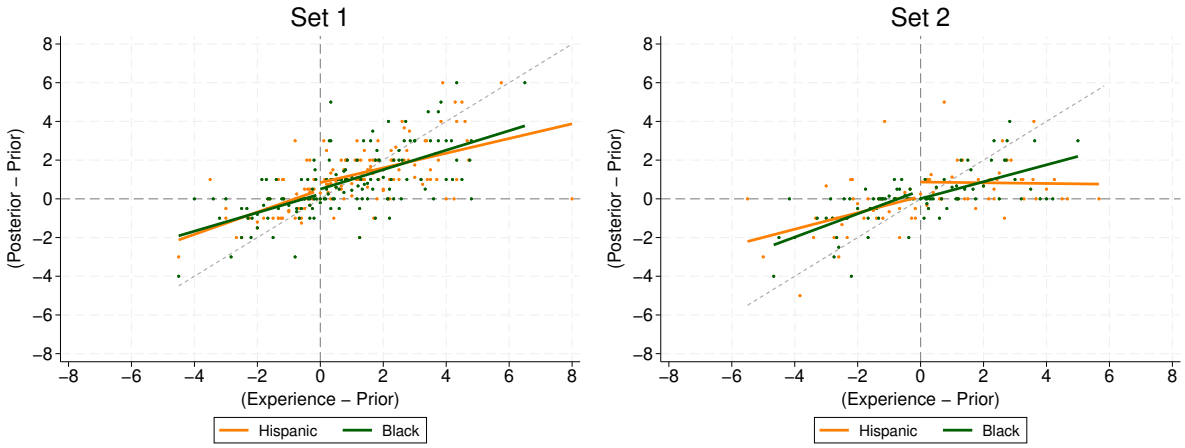
Figure B.1: Initial Beliefs about Workers' Performance
Biased Employers



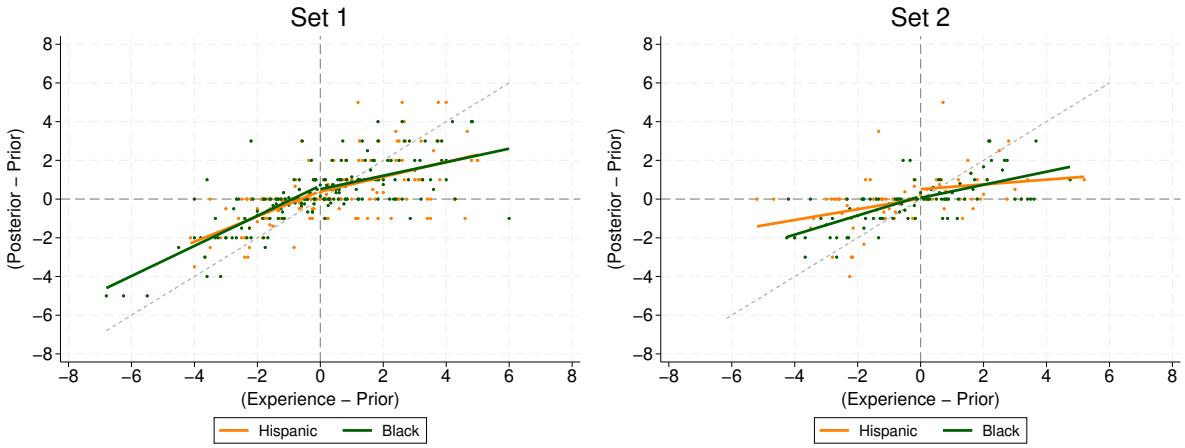
Notes: This figure shows the average initial beliefs stated by employers about each of the worker groups for the subsample of employers biased against both minority groups (45% of the sample). The first belief elicitation takes place before the first set of hiring decisions after employers are informed about the hiring task, but before they know which groups are available for hire. Confidence intervals included at the 95% level.

Figure B.2: Belief Updating
Heterogeneity by Initial Bias

(a) Biased Employers



(b) Unbiased Employers



Notes: These figures compare the difference in employer's beliefs after hiring decisions (posterior beliefs minus prior beliefs) and the hiring surprise (average hiring experience minus prior beliefs), averaged over bins. *Prior* is the employer's belief about the minority group's average performance before the hiring decisions. *Experience* is the average performance of the hired workers. *Posterior* is the employer's belief about the minority group's average performance after the set of hiring decisions. Panel A presents results for the subsample of employers biased against both minority groups (45% of the sample). Panel B presents results for the unbiased employers (remaining 55% of the sample). The left panels show statistics for the full sample in Set 1: treatments HH and HB for Hispanic workers and treatments BB and BH for Black workers. The right panels compare statistics for the second set of hiring decisions for the baseline treatments only: HH for Hispanic workers and BB for Black workers. The slope is based on a linear regression with a constant term.

Table B.1: Worker Sample Characteristics

	All	White	Hispanic	Black	p-val W=H	p-val W=B	p-val H=B
Panel A: Demographic Characteristics							
Age	28.37	28.87	26.62	28.62	0.001***	0.699	0.018**
	(4.04)	(3.84)	(4.22)	(4.07)			
Female	0.42	0.47	0.42	0.30	0.568	0.039**	0.215
	(0.50)	(0.50)	(0.50)	(0.46)			
Currently employed	0.18	0.19	0.18	0.16	0.836	0.601	0.793
	(0.39)	(0.40)	(0.39)	(0.37)			
Bachelors or higher	0.99	0.99	1.00	0.98	0.414	0.739	0.320
	(0.11)	(0.12)	(0.00)	(0.14)			
Panel B: Performance in the task							
Number of problems solved	5.01	5.18	4.76	4.74	0.227	0.218	0.964
	(2.17)	(2.13)	(2.10)	(2.34)			
Enjoy the task (1-10)	5.29	5.15	5.12	5.86	0.944	0.133	0.226
	(2.92)	(2.83)	(3.09)	(2.98)			
Difficulty of the task (1-10)	6.61	6.57	6.82	6.52	0.387	0.873	0.444
	(1.82)	(1.73)	(1.96)	(1.95)			
Which quartile you think you are (1-4)	2.49	2.43	2.66	2.52	0.135	0.561	0.510
	(0.99)	(0.94)	(1.00)	(1.11)			
Observations	250	150	50	50			

Notes: A total of 250 individuals participated in the online experiment as workers. Subjects were recruited from Prolific. Panel A reports the mean of each demographic variable across the different worker samples. Panel B reports the performance in the mathematical task. The corresponding standard deviation is reported in parentheses. *Enjoy the task* is the response to the question "From 1 to 10, how much did you enjoy solving the questions?." *Difficulty of the task* is the response to the question "Rate the difficulty of the questions, where 1 is extremely easy and 10 is extremely difficult." *Quartile* is the response to the question "Select how you think your performance compares to the performance of all the other Prolific workers who participated in this study", where 1 is the quartile 1 (top 25% of the workers), and 4 is the quartile 4 (bottom 25% of the workers). The last three columns report p-values from t-tests. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table B.2: Initial Priors About Worker Performance

	Initial Priors
Hispanic	-0.951***
	(0.038)
Black	-0.871***
	(0.040)
Constant	5.204***
	(0.035)
Observations	6,000
Number of employers	2,000
H_0: Hispanic = Black	0.007***

Notes: This table reports OLS estimates from a panel with 4,800 observations, corresponding to three observations per subject. Each employer reports three initial beliefs, one per worker group. The dependent variable is the initial belief about average group performance. *Hispanic* is a dummy equal to 1 if the initial prior is about Hispanic workers' performance. *Black* is a dummy equal to 1 if the initial prior is about Black workers' performance. The constant represents the initial prior about White workers' performance. Standard errors clustered at the individual level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table B.3: Hiring of Minority Workers (Set 1)

	Pr. Hiring Minority		
	(1)	(2)	(3)
Hispanic	-0.176***		-0.167***
	(0.026)		(0.021)
Black		-0.141***	-0.147***
		(0.021)	(0.018)
Constant	0.632***	0.596***	0.610***
	(0.017)	(0.014)	(0.011)
Observations	16,000	24,000	40,000
Number of employers	800	1,200	2,000
Time FE	Yes	Yes	Yes
Employer FE	Yes	Yes	Yes
H_0: Hispanic = Black	–	–	0.326

Notes: This table reports OLS estimates from two panels with 16,000 observations (columns 1 and 2), and one joint panel with 32,000 observations (column 3), corresponding to 20 observations per employer. The dependent variable is the probability that a worker is hired. *Hispanic* is a dummy variable equal to 1 if the hired worker was Hispanic. *Black* is a dummy variable equal to 1 if the hired worker was Black. Column 1 presents the estimates for hiring of Hispanic workers, pooling treatments HH and HB, when employers choose to hire between White and Hispanic workers in Set 1. Column 2 presents the estimates for hiring of Black workers, pooling treatments BB and BH, when employers choose to hire between White and Black workers in Set 1. Column 3 presents the estimates for Hispanic and Black workers, pooling all treatments in Set 1. The estimations include time fixed-effects. Robust standard errors clustered at the individual level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table B.4: Belief Updating - Differential Impact by Minority

	Posteriors about Minority	
	Set 1	Set 2
	(1)	(2)
Panel A: Belief Updating		
Prior	0.471***	0.594***
	(0.026)	(0.037)
Experience	0.573***	0.416***
	(0.027)	(0.038)
Prior × Hispanic	0.025	0.105*
	(0.044)	(0.056)
Experience × Hispanic	-0.013	-0.092
	(0.046)	(0.059)
Black Prior Mean	4.45	4.80
Black Posterior Mean	4.78	4.78
Observations	1,830	724
H_0: Prior = Experience	0.052*	0.016**
H_0: Prior × Hisp = Experience × Hisp	0.667	0.085*
Panel B: Test for Asymmetric Updating		
Prior	1.070***	1.027***
	(0.012)	(0.015)
Experience - Prior	0.515***	0.360***
	(0.040)	(0.065)
(Experience - Prior) × Negative Surprise	0.182**	0.123
	(0.074)	(0.104)
Prior × Hispanic	0.005	0.029
	(0.018)	(0.024)
(Experience - Prior) × Hispanic	0.007	-0.130
	(0.067)	(0.090)
(Experience - Prior) × Negative Surprise × Hispanic	-0.035	0.104
	(0.122)	(0.159)
Black Prior Mean	4.45	4.80
Black Posterior Mean	4.78	4.78
Observations	1,830	724
H_0: (Exp - Prior) × Neg + (Exp - Prior) × Neg × Hisp = 0	0.131	0.060*

Notes: This table reports OLS estimates. The dependent variables are posterior beliefs about average group performance (beliefs elicited after each set of hiring decisions). *Prior* is the employer's belief about the minority group's average performance before the hiring set. *Experience* is the average performance of the hired minority workers in each set. *Negative Surprise* is a dummy equal to 1 if the average hiring experience is less than the employer's priors about the average group productivity. *Hispanic* is a dummy if the employer was assigned to the HH or HB treatments, hence updating beliefs about Hispanic workers' performance. The estimations on Set 1 pool all treatments. Estimations on Set 2 use only the baseline treatments' data. This analysis restricts the sample to employers who hired at least one minority worker in the first set of hiring decisions, corresponding to 1,488 employers (93% of the sample). Robust standard errors are in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table B.5: Hiring Dynamics (Set 1) - Differential Impact by Minority

	Minority Hiring	
	(1)	(2)
Experience$_t$	0.064***	
	(0.004)	
Experience$_t$ × Hispanic	-0.003	
	(0.007)	
Positive Surprise$_t$		0.165***
		(0.031)
Negative Surprise$_t$		-0.147***
		(0.032)
Positive Surprise$_t$ × Hispanic		-0.065
		(0.050)
Negative Surprise$_t$ × Hispanic		-0.017
		(0.052)
Hispanic	0.073	0.096
	(0.066)	(0.066)
Prior about Black	0.062***	0.076***
	(0.022)	(0.022)
Prior about Hispanic	0.016	0.042*
	(0.022)	(0.023)
Prior about White	-0.070***	-0.068***
	(0.019)	(0.019)
Black Control Mean	2.247	2.247
Observations	8,935	8,935
Number of employers	1,874	1,874
Time FE	Yes	Yes
H_0: Pos + Neg = 0		0.756
H_0: Pos × Hisp + Neg × Hisp = 0		0.390

Notes: This table reports OLS estimates of the impact of hiring experiences with minority workers on subsequent hiring of the group, i.e., the number of minority workers hired in the following periods, holding constant the number of minority workers hired across employers. *Experience* is the number of problems solved correctly by a hired minority worker. *Positive Surprise* is a dummy variable equal to 1 if the minority worker's performance is greater than the employer's priors about average performance. *Negative Surprise* is a dummy variable equal to 1 if the minority worker's performance is less than the employer's priors about average performance. *Hispanic* is a dummy if the employer was assigned to the HH or HB treatments, hence subsequent hiring of Hispanic workers. Estimations include all the treatments. The estimations include time fixed-effects. Robust standard errors clustered at the individual level are in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table B.6: Robustness: Employer Priors in the Placebo Experiment

	Race	Neutral	p-val
Panel A: Initial bias			
Initial bias against Hispanic	-1.00	-0.73	0.005***
	(1.63)	(2.02)	
Initial bias against Black	-0.95	-0.57	0.000***
	(1.66)	(2.29)	
Panel B: Distribution of employers			
% Biased against Both	0.45	0.39	0.027**
	(0.50)	(0.49)	
% Correct beliefs	0.30	0.23	0.007***
	(0.46)	(0.42)	
% Biased in favor Both	0.05	0.17	0.000***
	(0.21)	(0.38)	
Observations	1,600	400	

Notes: A total of 1,600 individuals participated as employers in the main experiment (Race), and 400 individuals participated as employers in the placebo experiment (Neutral). Subjects were recruited from Prolific. Panel A reports the mean initial bias against minority workers, defined as the difference between the priors about minority workers' performance and priors about White workers' performance. The corresponding standard deviation is reported in parentheses. Panel B reports the distribution of employers. *% Biased against Both* is the percentage of employers biased against both Black and Hispanic workers. *% Correct beliefs* is the percentage of employers who stated 5 as the average performance of the workers from the different groups. *% Biased in favor Both* is the percentage of employers biased in favor of both Black and Hispanic workers. The last column reports p-values from a t-test against the null hypothesis that subjects across the main experiment and the placebo experiment are different from each other. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table B.7: Belief Updating - Placebo Experiment

	Posteriors about Minority		Posteriors about Majority	
	Set 1	Set 2	Set 1	Set 2
	(1)	(2)	(3)	(4)
Panel A: Belief Updating				
Prior	0.494***	0.699***	0.530***	0.728***
	(0.035)	(0.041)	(0.030)	(0.025)
Experience	0.561***	0.327***	0.497***	0.290***
	(0.036)	(0.042)	(0.032)	(0.027)
Prior × Neutral	-0.065	-0.180**	-0.221***	-0.161***
	(0.056)	(0.072)	(0.057)	(0.044)
Experience × Neutral	0.030	0.146**	0.234***	0.161***
	(0.058)	(0.071)	(0.059)	(0.048)
Race Prior Mean	4.22	4.62	5.25	5.37
Race Posterior Mean	4.69	4.69	5.38	5.38
Observations	1,124	554	1,172	1,147
H_0: Prior = Experience	0.348	0.000***	0.597	0.000***
H_0: Prior × Neutral = Experience × Neutral	0.400	0.021**	0.000***	0.000***

Notes: This table reports OLS estimates. The dependent variables are posterior beliefs about average group performance (beliefs elicited after each set of hiring decisions). *Prior* is the employer's belief about the minority group's average performance before the hiring set. *Experience* is the average performance of the hired minority workers in each set. *Neutral* is a dummy variable equal to 1 if the employer was assigned to the placebo experiment, using neutral labels instead of racial and ethnic groups. Columns 1 and 2 present the estimates for beliefs about minority workers' average performance, Hispanic in the main experiment and Orange in the placebo experiment. Columns 3 and 4 replicate the analysis for beliefs majority workers' average performance, White in the main experiment and Purple in the placebo experiment. Estimations pool treatments HH and HB for the main experiment and the treatments "neutral baseline" and "neutral spillover" in the placebo experiment, except for posteriors about minority workers' performance after Set 2, which only employ HH and "neutral baseline" treatments data. The analysis in columns 1 and 2 restricts the sample to employers who hired at least one minority worker in the first and second sets of hiring decisions. The analysis in columns 3 and 4 restricts the sample to employers who hired at least one majority worker in the first and second sets of hiring decisions. Robust standard errors are in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table B.8: Spillovers in Beliefs - Placebo Experiment

	Posteriors about Green
Priors about Green	0.980***
	(0.011)
Hiring Surprise with Orange	0.129***
	(0.037)
Priors about Green × Neutral	-0.085***
	(0.024)
Hiring Surprise with Orange × Neutral	-0.113
	(0.070)
Prior Mean	4.36
Posterior Mean	4.36
Observations	1,124
H_0: Surprise + Surprise × Neutral = 0	0.784

Notes: This table reports OLS estimates. The dependent variable is the posterior beliefs about Green (Black) workers' performance after the first set of hiring decisions in treatments HH, HB, and "neutral baseline" and "neutral spillover". *Priors about Green* is the employer's beliefs about Green (Black) workers' average performance before Set 1. *Hiring Surprise with Orange* is the difference between the average performance of the hired Hispanic (Orange) workers and the prior beliefs about Hispanic (Orange) workers' average performance. *Neutral* is a dummy variable equal to 1 if the employer was assigned to the placebo experiment, using neutral labels instead of racial and ethnic groups. The analysis restricts the sample to employers who hired at least one Hispanic (Orange) worker in Set 1. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table B.9: Robustness Checks: Guessing the Purpose of the Study and Inattention

	All	HH	BB	HB	BH	ANOVA p-val	Baseline Treat.	Spillovers Treat.	t-test p-val
Panel A: Experimenter Demand Effect									
Discrimination/Racism	0.03	0.04	0.02	0.04	0.03	0.559	0.03	0.04	0.398
Prejudice/Stereotypes	0.09	0.11	0.08	0.10	0.08	0.424	0.10	0.09	0.730
Any of the above	0.12	0.14	0.10	0.14	0.12	0.236	0.12	0.13	0.880
Panel B: Inattention									
Failed attention check	0.07	0.09	0.05	0.07	0.07	0.265	0.07	0.07	0.766
Observations	1,600	400	400	400	400	–	800	800	–

Notes: Panel A reports the percentage of subjects who are able to guess the purpose of the experiment correctly, as seen by the answer to the question: *"What do you think this was about?"*, at the end of the experiment. The first row reports the percentage of subjects whose response is related to discrimination. The second row reports the percentage of subjects whose response is related to bias, prejudice, or stereotypes. The third row aggregates the results of the first two rows. Panel B reports the percentage of subjects who failed one attention check.

Table B.10: Robustness Checks: Belief Updating

	Excluding subjects who guessed the experiment's purpose				Excluding subjects who failed one attention check			
	Posteriors about Hispanic		Posteriors about Black		Posteriors about Hispanic		Posteriors about Black	
	Set 1	Set 2	Set 1	Set 2	Set 1	Set 2	Set 1	Set 2
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Belief Updating								
Prior	0.506***	0.703***	0.483***	0.601***	0.479***	0.674***	0.478***	0.594***
	(0.038)	(0.046)	(0.034)	(0.039)	(0.035)	(0.044)	(0.033)	(0.038)
Experience	0.554***	0.315***	0.572***	0.414***	0.579***	0.348***	0.574***	0.414***
	(0.039)	(0.048)	(0.035)	(0.041)	(0.036)	(0.046)	(0.034)	(0.040)
Prior Mean	4.24	4.71	4.47	4.81	4.29	4.76	4.46	4.80
Posterior Mean	4.72	4.73	4.82	4.80	4.75	4.77	4.80	4.78
Observations	626	305	647	328	670	325	687	347
Treatments	HH, HB	HH	BB, BH	BB	HH, HB	HH	BB, BH	BB
H_0: Prior = Experience	0.535	0.000***	0.198	0.020**	0.155	0.000***	0.154	0.020**
Panel B: Test for Asymmetric Updating								
Prior	1.082***	1.050***	1.073***	1.031***	1.070***	1.056***	1.062***	1.028***
	(0.016)	(0.020)	(0.015)	(0.016)	(0.014)	(0.019)	(0.014)	(0.014)
Experience - Prior	0.508***	0.224***	0.527***	0.357***	0.552***	0.246***	0.548***	0.351***
	(0.057)	(0.064)	(0.056)	(0.072)	(0.054)	(0.069)	(0.054)	(0.066)
(Experience - Prior) × Negative Surprise	0.174*	0.223*	0.123	0.122	0.095	0.238*	0.071	0.139
	(0.105)	(0.126)	(0.114)	(0.114)	(0.097)	(0.125)	(0.106)	(0.106)
Prior Mean	4.24	4.71	4.47	4.81	4.29	4.76	4.46	4.80
Posterior Mean	4.72	4.73	4.82	4.80	4.75	4.77	4.80	4.78
Observations	626	305	647	328	670	325	687	347
Treatments	HH, HB	HH	BB, BH	BB	HH, HB	HH	BB, BH	BB

Notes: This table reports OLS estimates. The dependent variables are posterior beliefs about average group performance (beliefs elicited after each set of hiring decisions). The left panel excluded the 12% of subjects who correctly guessed the experiment's purpose. The right panel excluded the 7% of subjects who failed one attention check. *Prior* is the employer's belief about the minority group's average performance before the hiring set. *Experience* is the average performance of the hired minority workers in each set. *Negative Surprise* is a dummy equal to 1 if the average hiring experience is less than the employer's priors about the average group performance. The estimations on Set 1 pool treatments HH and HB for posteriors about Hispanic workers' performance and treatments BB and BH for posteriors about Black workers' performance. Estimations on Set 2 use only the baseline treatments' data. This analysis restricts the sample to employers who hired at least one minority worker in both sets of hiring decisions. Robust standard errors are in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table B.11: Robustness Checks: Spillovers in Beliefs

	Excluding subjects who guessed purpose		Excluding subjects who failed attention	
	Posteriors about:		Posteriors about:	
	Hispanic	Black	Hispanic	Black
	(1)	(2)	(1)	(2)
Panel A: Belief Updating				
Prior	0.960***	0.977***	0.961***	0.986***
	(0.012)	(0.012)	(0.011)	(0.011)
Hiring Surprise with Other Minority	0.177***	0.123***	0.166***	0.128***
	(0.034)	(0.039)	(0.034)	(0.038)
Prior Mean	4.31	4.28	4.34	4.32
Posterior Mean	4.28	4.37	4.32	4.44
Observations	666	634	706	680
Treatments	BB, BH	HH, HB	BB, BH	HH, HB
Panel B: Test for Asymmetric Updating				
Prior	0.951***	0.951***	0.949***	0.950***
	(0.015)	(0.018)	(0.014)	(0.016)
Hiring Surprise with Other Minority	0.202***	0.183***	0.199***	0.220***
	(0.044)	(0.055)	(0.045)	(0.054)
Hiring Surprise with Other Minority × Neg. Surp.	-0.063	-0.209*	-0.087	-0.297**
	(0.095)	(0.125)	(0.093)	(0.120)
Prior Mean	4.31	4.28	4.34	4.32
Posterior Mean	4.28	4.37	4.32	4.44
Observations	666	634	706	680
Treatments	BB, BH	HH, HB	BB, BH	HH, HB
H_0: Surp + Surp × NegSurp = 0	0.062*	0.790	0.118	0.396

Notes: This table reports OLS estimates. The dependent variable is the posterior beliefs after the first set of hiring decisions about Hispanic workers' performance (column 1) and about Black workers' performance (column 2). The left panel excluded the 12% of subjects who correctly guessed the experiment's purpose. The right panel excluded the 7% of subjects who failed one attention check. *Prior* is the employer's beliefs about minority workers' performance before the first hiring set. *Hiring Surprise with Other Minority* is the difference between the average performance of the hired Black workers and the prior beliefs about Black workers' performance (column 1), and the difference between the average performance of the hired Hispanic workers and the prior beliefs about Hispanic workers' performance (column 2). *Negative Surprise* is a dummy equal to 1 if the hiring surprise is negative. The estimations on beliefs about Hispanic workers' performance pool treatments BB and BH, when employers choose to hire between White and Black workers in Set 1. The estimations on beliefs about Black workers' performance pool treatments HH and HB, when employers choose to hire between White and Hispanic workers in Set 1. The analysis restricts the sample to employers who hired at least one Black or one Hispanic worker in Set 1. Robust standard errors are in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table B.12: Robustness Checks: Hiring Dynamics (Set 1)

	Excluding subjects who guessed the experiment's purpose				Excluding subjects who failed one attention check			
	Hispanic Hiring		Black Hiring		Hispanic Hiring		Black Hiring	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Experience$_t$	0.060***		0.060***		0.069***		0.061***	
	(0.006)		(0.005)		(0.006)		(0.005)	
Positive Surprise$_t$		0.096**		0.154***		0.149***		0.136***
		(0.044)		(0.039)		(0.040)		(0.040)
Negative Surprise$_t$		-0.184***		-0.152***		-0.164***		-0.176***
		(0.045)		(0.041)		(0.044)		(0.041)
Prior about Minority	0.082**	0.125***	0.019	0.058	0.086**	0.136***	0.030	0.069**
	(0.035)	(0.035)	(0.036)	(0.036)	(0.036)	(0.036)	(0.034)	(0.035)
Prior about White	-0.050	-0.049	-0.046	-0.046	-0.059*	-0.058*	-0.068**	-0.068**
	(0.033)	(0.033)	(0.034)	(0.035)	(0.034)	(0.034)	(0.032)	(0.032)
Control Mean	2.268	2.268	2.222	2.222	2.244	2.244	2.238	2.238
Observations	3,035	3,035	3,184	3,184	3,221	3,221	3,395	3,395
Number of employers	634	634	666	666	680	680	706	706
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
H_0: Pos + Neg = 0		0.286		0.987		0.849		0.597

Notes: This table reports OLS estimates of the impact of hiring experiences with minority workers on subsequent hiring of the group, i.e., the number of minority workers hired in the following periods, holding constant the number of minority workers hired across employers. The left panel excluded the 12% of subjects who correctly guessed the experiment's purpose. The right panel excluded the 7% of subjects who failed one attention check. *Experience* is the number of problems solved by a hired minority worker. *Positive Surprise* is a dummy variable equal to 1 if the minority worker's performance is greater than the employer's priors about the average performance. *Negative Surprise* is a dummy variable equal to 1 if the minority worker's performance is less than the employer's priors about the average performance. Columns 1, 2, 5, and 6 present the estimates for hiring of Hispanic workers, pooling treatments HH and HB, when employers choose to hire between White and Hispanic workers in Set 1. Columns 3, 4, 7, and 8 present the estimates for hiring of Black workers, pooling treatments BB and BH, when employers choose to hire between White and Black workers in Set 1. The estimations include time fixed-effects. Robust standard errors clustered at the individual level are in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table B.13: Robustness Checks: Spillovers in Hiring

	Excluding subjects who guessed purpose		Excluding subjects who failed attention	
	Pr. Hiring		Pr. Hiring	
	Hispanic	Black	Hispanic	Black
	(1)	(2)	(3)	(4)
Set 2	0.018*	-0.002	0.014	-0.004
	(0.011)	(0.010)	(0.010)	(0.010)
Spillover Tr.	0.003	0.002	0.009	-0.010
	(0.016)	(0.015)	(0.016)	(0.015)
Set 2 × Spillover Tr.	0.023	0.047***	0.024	0.055***
	(0.016)	(0.016)	(0.015)	(0.015)
Priors about Hispanic	0.015*	0.018***	0.017**	0.019***
	(0.008)	(0.006)	(0.007)	(0.006)
Priors about Black	0.023***	0.029***	0.026***	0.032***
	(0.007)	(0.007)	(0.006)	(0.006)
Prior about White	-0.025***	-0.027***	-0.034***	-0.030***
	(0.005)	(0.006)	(0.005)	(0.005)
Control Mean	0.429	0.459	0.434	0.454
Observations	13,920	14,080	14,780	15,040
Number of employers	696	704	739	752
Period FE	Yes	Yes	Yes	Yes

Notes: This table reports OLS estimates using the panel of 20 hiring decisions. The dependent variable is a dummy equal to 1 if a minority worker was hired in a given time period. The left panel excluded the 12% of subjects who correctly guessed the experiment's purpose. The right panel excluded the 7% of subjects who failed one attention check. *Set2* is a dummy equal to 1 for the second set of hiring decisions, i.e., periods 11 to 20. *Spillover Tr.* is a dummy variable equal to 1 if the employer was assigned to the spillover treatments: HB for Black hiring and BH for Hispanic hiring. Columns 1 and 3 present the estimates for Hispanic hiring, pooling treatments HH and BH, when employers choose to hire between White and Hispanic workers in Set 2. Columns 2 and 4 present the estimates for Black hiring, pooling treatments BB and HB, when employers choose to hire between White and Hispanic workers in Set 2. The estimations include time (within set) fixed-effects. Robust standard errors clustered at the individual level are in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table B.14: Heterogeneity by Social Desirability Bias: Initial Priors

	Initial Priors
Hispanic	-0.984***
	(0.055)
Black	-0.953***
	(0.056)
Hispanic × High SD	-0.045
	(0.082)
Black × High SD	0.013
	(0.083)
High SD	0.011
	(0.076)
Constant	5.241***
	(0.049)
Observations	4,800
Number of employers	1,600
H_0: Hispanic = Black	0.428

Notes: This table reports OLS estimates from a panel with 4,800 observations, corresponding to three observations per subject. Each employer reports three initial beliefs, one per worker group. The dependent variable is the initial belief about average group productivity. The Social Desirability index is an index ranging from 0 to 1 that measures a subject's propensity to give socially desirable responses. I used the 7-item X1 scale as described in Fisher (1993) and constructed the social desirability index using Anderson (2008)'s method. Section C.3.1 presents the questions used. *Hispanic* is a dummy equal to 1 if the initial prior is about Hispanic workers' performance. *Black* is a dummy equal to 1 if the initial prior is about Black workers' performance. *High SD* is a dummy equal to 1 if the subject has an above-median social desirability index. The constant represents the initial prior about White workers' performance. Robust standard errors clustered at the individual level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table B.15: Heterogeneity by Social Desirability Bias: Hiring of Minority Workers

	Pr. Hiring Minority		
	(1)	(2)	(3)
Hispanic	-0.188***		-0.169***
	(0.031)		(0.021)
Hispanic × High SD	0.025		0.018
	(0.032)		(0.023)
Black		-0.141***	-0.154***
		(0.024)	(0.019)
Black × High SD		0.001	0.006
		(0.024)	(0.021)
High SD	-0.012	-0.001	-0.005
	(0.016)	(0.012)	(0.010)
Constant	0.638***	0.596***	0.613***
	(0.019)	(0.015)	(0.012)
Observations	16,000	24,000	40,000
Number of employers	800	1,200	2,000
Time FE	Yes	Yes	Yes
H_0: Hispanic = Black	–	–	0.267

Notes: This table reports OLS estimates from two panels with 16,000 observations (columns 1 and 2), and one joint panel with 32,000 observations (column 3), corresponding to 20 observations per employer. The dependent variable is the probability that a worker is hired. The Social Desirability index is an index ranging from 0 to 1 that measures a subject's propensity to give socially desirable responses. I used the 7-item X1 scale as described in Fisher (1993) and constructed the social desirability index using Anderson (2008)'s method. Section C.3.1 presents the questions used. *Hispanic* is a dummy variable equal to 1 if the hired worker was Hispanic. *Black* is a dummy variable equal to 1 if the hired worker was Black. *High SD* is a dummy equal to 1 if the subject has an above-median social desirability index. Column 1 presents the estimates for hiring of Hispanic workers, pooling treatments HH and HB, when employers choose to hire between White and Hispanic workers in Set 1. Column 2 presents the estimates for hiring of Black workers, pooling treatments BB and BH, when employers choose to hire between White and Black workers in Set 1. Column 3 presents the estimates for Hispanic and Black workers, pooling all treatments in Set 1. Robust standard errors clustered at the individual level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table B.16: Belief Updating about White Workers

	Posteriors about White	
	Set 1	Set 2
	(1)	(2)
Panel A: Belief Updating		
Prior	0.476***	0.681***
	(0.021)	(0.016)
Experience	0.556***	0.333***
	(0.022)	(0.017)
Prior Mean	5.22	5.30
Posterior Mean	5.30	5.29
Observations	1,910	1,910
Panel B: Test for Asymmetric Updating		
Prior	1.006***	1.021***
	(0.009)	(0.008)
Experience - Prior	0.596***	0.319***
	(0.029)	(0.023)
(Experience - Prior) × Neg. Surp.	0.281***	-0.087
	(0.081)	(0.071)
Prior Mean	5.22	5.30
Posterior Mean	5.30	5.29
Observations	1,910	1,910

Notes: This table reports OLS estimates. The dependent variables are posterior beliefs about average White performance (beliefs elicited after each set of hiring decisions). *Prior* is the employer's belief about White workers' average performance before the hiring set. *Experience* is the average performance of the hired White workers in each set. *Negative Surprise* is a dummy equal to 1 if the average hiring experience is less than the employer's priors about the average White performance. The estimations pool all the treatments. This analysis restricts the sample to employers who hired at least one White worker in both sets of hiring decisions, corresponding to 1,531 employers (96% of the sample). Robust standard errors are in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table B.17: Hiring Dynamics for White Workers (Set 1)

	White Hiring	
	(1)	(2)
Experience$_t$	0.046***	
	(0.003)	
Positive Surprise$_t$		0.093***
		(0.021)
Negative Surprise$_t$		-0.118***
		(0.021)
Prior about Minority	-0.138***	-0.139***
	(0.020)	(0.020)
Prior about White	0.087***	0.118***
	(0.020)	(0.020)
Control Mean	2.688	2.688
Observations	11,065	11,065
Clusters (Employers)	1,957	1,957
Time FE	Yes	Yes
H_0: Pos + Neg = 0	–	0.515

Notes: This table reports OLS estimates of the impact of hiring experiences with White workers on subsequent hiring of White workers, i.e., the number of White workers hired in the following periods, holding constant the number of White workers hired across employers. *Experience* is the number of problems solved correctly by a hired White worker. *Positive Surprise* is a dummy variable equal to 1 if the White worker's performance is greater than the employer's priors about the average performance. *Negative Surprise* is a dummy variable equal to 1 if the White worker's performance is less than the employer's priors about the average performance. Estimations pool all treatments. The estimations include time fixed-effects. Robust standard errors clustered at the individual level are in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table B.18: Belief Updating
Heterogeneity by Initial Bias

	Biased Employers				Unbiased Employers			
	Posteriors about Hispanic		Posteriors about Black		Posteriors about Hispanic		Posteriors about Black	
	Set 1	Set 2	Set 1	Set 2	Set 1	Set 2	Set 1	Set 2
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Belief Updating								
Prior	0.515***	0.696***	0.472***	0.588***	0.508***	0.705***	0.509***	0.603***
	(0.061)	(0.065)	(0.051)	(0.055)	(0.045)	(0.055)	(0.043)	(0.049)
Experience	0.570***	0.329***	0.592***	0.433***	0.530***	0.318***	0.538***	0.399***
	(0.057)	(0.065)	(0.049)	(0.056)	(0.049)	(0.059)	(0.046)	(0.052)
Prior Mean	3.84	4.48	4.06	4.65	4.65	4.96	4.72	4.91
Posterior Mean	4.62	4.59	4.72	4.70	4.84	4.94	4.87	4.84
Observations	329	164	304	162	398	193	427	205
Treatments	HH, HB	HH	BB, BH	BB	HH, HB	HH	BB, BH	BB
H_0: Prior = Experience	0.640	0.005***	0.225	0.160	0.812	0.001***	0.738	0.043**
Panel B: Test for Asymmetric Updating								
Prior	1.101***	1.078***	1.046***	1.049***	1.058***	1.037***	1.080***	1.012***
	(0.024)	(0.031)	(0.025)	(0.027)	(0.017)	(0.021)	(0.015)	(0.017)
Experience - Prior	0.539***	0.208**	0.626***	0.349***	0.479***	0.258***	0.431***	0.364***
	(0.077)	(0.082)	(0.071)	(0.095)	(0.074)	(0.091)	(0.073)	(0.090)
(Experience - Prior) × Negative Surprise	0.148	0.336*	-0.133	0.192	0.159	0.119	0.248**	0.076
	(0.163)	(0.189)	(0.158)	(0.158)	(0.120)	(0.146)	(0.139)	(0.139)
Prior Mean	3.84	4.48	4.06	4.65	4.65	4.96	4.72	4.91
Posterior Mean	4.62	4.59	4.72	4.70	4.84	4.94	4.87	4.84
Observations	329	164	304	162	398	193	427	205
Treatments	HH, HB	HH	BB, BH	BB	HH, HB	HH	BB, BH	BB

Notes: This table reports OLS estimates. The dependent variables are posterior beliefs about average gorup performance (beliefs elicited after each set of hiring decisions). The left panel limits the analysis to biased subjects (45% of the sample). The right panel limits the analysis to unbiased subjects (the remaining 55% of the sample). *Prior* is the employer's belief about the minority group's average performance before the hiring set. *Experience* is the average performance of the hired minority workers in each set. *Negative Surprise* is a dummy equal to 1 if the average hiring experience is less than the employer's priors about the average group performance. The estimations on Set 1 pool treatments HH and HB for posteriors about Hispanic workers' performance and treatments BB and BH for posteriors about Black workers' performance. Estimations on Set 2 use only the baseline treatments' data. This analysis restricts the sample to employers who hired at least one minority worker in both sets of hiring decisions. Robust standard errors are in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table B.19: Spillovers in Beliefs
Heterogeneity by Initial Bias

	Biased Employers		Unbiased Employers	
	Posteriors about:		Posteriors about:	
	Hispanic	Black	Hispanic	Black
	(1)	(2)	(1)	(2)
Panel A: Belief Updating				
Prior	0.966***	0.995***	0.960***	0.972***
	(0.019)	(0.018)	(0.013)	(0.015)
Hiring Surprise with Other Minority	0.155***	0.121**	0.172***	0.125**
	(0.045)	(0.055)	(0.046)	(0.053)
Prior Mean	3.81	3.77	4.68	4.76
Posterior Mean	3.89	4.01	4.58	4.74
Observations	313	336	437	402
Treatments	BB, BH	HH, HB	BB, BH	HH, HB
Panel B: Test for Asymmetric Updating				
Prior	0.977***	0.965***	0.939***	0.946***
	(0.024)	(0.027)	(0.017)	(0.021)
Hiring Surprise with Other Minority	0.133***	0.184**	0.251***	0.201**
	(0.049)	(0.072)	(0.072)	(0.080)
Hiring Surprise with Other Minority × Neg. Surp.	0.080	-0.287	-0.171	-0.215
	(0.160)	(0.203)	(0.121)	(0.147)
Prior Mean	3.81	3.77	4.68	4.76
Posterior Mean	3.89	4.01	4.58	4.74
Observations	313	336	437	402
Treatments	BB, BH	HH, HB	BB, BH	HH, HB
H_0: Surp + Surp × NegSurp = 0	0.128	0.541	0.321	0.887

Notes: This table reports OLS estimates. The dependent variable is the posterior beliefs after the first set of hiring decisions about Hispanic workers' performance (column 1) and about Black workers' performance (column 2). The left panel limits the analysis to biased subjects (45% of the sample). The right panel limits the analysis to unbiased subjects (the remaining 55% of the sample). *Prior* is the employer's beliefs about minority workers' performance before the first hiring set. *Hiring Surprise with Other Minority* is the difference between the average performance of the hired Black workers and the prior beliefs about Black workers' performance (column 1), and the difference between the average performance of the hired Hispanic workers and the prior beliefs about Hispanic workers' performance (column 2). *Negative Surprise* is a dummy equal to 1 if the hiring surprise is negative. The estimations on beliefs about Hispanic workers' performance pool treatments BB and BH, when employers choose to hire between White and Black workers in Set 1. The estimations on beliefs about Black workers' performance pool treatments HH and HB, when employers choose to hire between White and Hispanic workers in Set 1. The analysis restricts the sample to employers who hired at least one Black or one Hispanic worker in Set 1. Robust standard errors are in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table B.20: Hiring Dynamics (Set 1)
Heterogeneity by Initial Bias

	Biased Employers				Unbiased Employers			
	Hispanic Hiring		Black Hiring		Hispanic Hiring		Black Hiring	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Experience$_t$	0.087***		0.058***		0.049***		0.064***	
	(0.010)		(0.008)		(0.007)		(0.006)	
Positive Surprise$_t$		0.205***		0.152**		0.033		0.137***
		(0.060)		(0.063)		(0.051)		(0.049)
Negative Surprise$_t$		-0.197***		-0.182***		-0.180***		-0.156***
		(0.063)		(0.069)		(0.052)		(0.049)
Prior about Minority	-0.077	-0.014	-0.020	0.026	0.075	0.105*	0.014	0.050
	(0.059)	(0.061)	(0.066)	(0.066)	(0.060)	(0.060)	(0.061)	(0.061)
Prior about White	0.112**	0.114**	-0.001	-0.003	-0.043	-0.040	-0.044	-0.043
	(0.054)	(0.055)	(0.065)	(0.065)	(0.072)	(0.072)	(0.065)	(0.066)
Control Mean	1.991	1.991	2.059	2.059	2.455	2.455	2.366	2.366
Observations	1,419	1,419	1,375	1,375	2,108	2,108	2,243	2,243
Number of employers	336	336	313	313	402	402	437	437
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
H_0: Pos + Neg = 0		0.944		0.814		0.124		0.839

Notes: This table reports OLS estimates of the impact of hiring experiences with minority workers on subsequent hiring of the group, i.e., the number of minority workers hired in the following periods, holding constant the number of minority workers hired across employers. The left panel presents the results for the biased employers (45% of the sample). The right panel presents the results for the unbiased employers (the remaining 55% of the sample). *Experience* is the number of problems solved correctly by a hired minority worker. *Positive Surprise* is a dummy variable equal to 1 if the minority worker's performance is greater than the employer's priors about the average performance. *Negative Surprise* is a dummy variable equal to 1 if the minority worker's performance is less than the employer's priors about the average performance. Columns 1, 2, 5, and 6 present the estimates for hiring of Hispanic workers, pooling treatments HH and HB, when employers choose to hire between White and Hispanic workers in Set 1. Columns 3, 4, 7, and 8 present the estimates for hiring of Black workers, pooling treatments BB and BH, when employers choose to hire between White and Black workers in Set 1. The estimations include time fixed-effects. Robust standard errors clustered at the individual level are in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table B.21: Spillovers in Hiring
Heterogeneity by Initial Bias

	Biased Employers		Unbiased Employers	
	Pr. Hiring		Pr. Hiring	
	Hispanic	Black	Hispanic	Black
	(1)	(2)	(3)	(4)
Set 2	0.030**	0.010	0.002	-0.013
	(0.014)	(0.013)	(0.013)	(0.014)
Spillover Tr.	0.023	-0.013	-0.013	0.001
	(0.024)	(0.020)	(0.020)	(0.020)
Set 2 × Spillover Tr.	0.034	0.060***	0.016	0.044**
	(0.022)	(0.021)	(0.018)	(0.020)
Priors about Hispanic	0.006	0.007	0.013	0.024**
	(0.012)	(0.009)	(0.011)	(0.011)
Priors about Black	0.035***	0.040***	0.002	0.015
	(0.010)	(0.010)	(0.010)	(0.010)
Prior about White	-0.023**	-0.027***	-0.009	-0.020
	(0.009)	(0.010)	(0.014)	(0.016)
Control Mean	0.371	0.398	0.490	0.506
Observations	7,220	7,300	8,780	8,700
Number of employers	361	365	439	435
Period FE	Yes	Yes	Yes	Yes

Notes: This table reports OLS estimates using the panel of 20 hiring decisions. The left panel presents the results for biased employers (45% of the sample). The right panel presents the results for unbiased employers (55% of the sample). The dependent variable is a dummy equal to 1 if a minority worker was hired in a given time period. *Set 2* is a dummy equal to 1 for the second set of hiring decisions, i.e., periods 11 to 20. *Spillover Tr.* is a dummy variable equal to 1 if the employer was assigned to the spillover treatments: HB for Black hiring and BH for Hispanic hiring. Columns 1 and 3 present the estimates for Hispanic hiring, pooling treatments HH and BH, when employers choose to hire between White and Hispanic workers in Set 2. Columns 2 and 4 present the estimates for Black hiring, pooling treatments BB and HB, when employers choose to hire between White and Hispanic workers in Set 2. The estimations include time (within set) fixed-effects. Robust standard errors clustered at the individual level are in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table B.22: Source of Discrimination

	Hispanic workers			Black workers		
	Diff. in hiring	Diff. in beliefs	Residual	Diff. in hiring	Diff. in beliefs	Residual
Panel A: Set 1						
Baseline	-1.183	-1.030	-0.152	-0.987	-0.787	-0.199
Panel B: Set 2						
Baseline	-1.085	-0.743	-0.343	-0.810	-1.137	-0.110
Spillovers	-0.410	-0.990	0.580	0.015	-0.797	0.812

Notes: This table reports the average difference in hiring, defined as the number of White workers hired minus the number of minority workers hired, and the average difference in beliefs, defined as the beliefs about White workers' performance minus the beliefs about minority workers' performance. The residual is calculated as the difference between the difference in hiring and the difference in beliefs.

Table B.23: Impact of Natural Exposure

	Priors about:			Pr. Hiring	
	White	Hispanic	Black	Hispanic	Black
% Hispanic population	-0.097	0.244	0.087	0.051	0.116**
	(0.265)	(0.274)	(0.280)	(0.058)	(0.049)
% Black population	0.101	-0.025	-0.038	-0.049	0.026
	(0.322)	(0.308)	(0.330)	(0.073)	(0.071)
Dep. Variable Mean	5.246	4.242	4.299	0.441	0.452
Observations	1,434	1,434	1,434	7,180	7,160
Number of employers	1,434	1,434	1,434	718	716
Time FE	–	–	–	Yes	Yes

Notes: This table reports OLS estimates. The left panel shows estimates for the prior beliefs about the average performance of each worker group independently. The right panel shows estimates for the probability of hiring a minority worker. *% Hispanic population* is the proportion of Hispanic individuals living in the employers' county of residence. *% Black population* is the proportion of Black individuals living in the employers' county of residence. The analysis is limited to employers whose location was matched with the proportion of Hispanic and Black individuals at the county level. Robust standard errors are in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table B.24: Threshold Regression: Posterior Beliefs About Hispanic

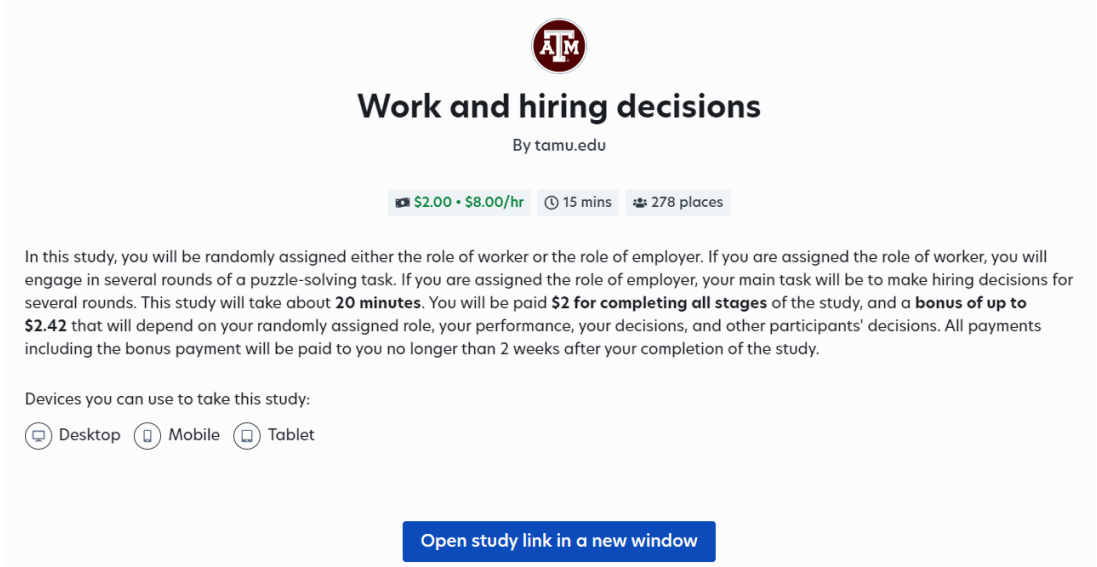
	% Hispanic Pop. > 0.293		% Hispanic Pop. < 0.293	
	Set 1	Set 2	Set 1	Set 2
	(1)	(2)	(3)	(4)
Prior	0.476***	0.540***	0.504***	0.762***
	(0.059)	(0.078)	(0.044)	(0.049)
Experience	0.583***	0.489***	0.552***	0.260***
	(0.059)	(0.081)	(0.045)	(0.052)
Prior Mean	4.30	4.67	4.28	4.77
Posterior Mean	4.70	4.67	4.76	4.82
Observations	196	107	531	250
Treatments	HH, HB	HH	HH, HB	HH
H_0: Prior = Experience	0.360	0.746	0.587	0.000***

Notes: This table reports OLS estimates. The dependent variables are posterior beliefs (beliefs elicited after each set of hiring decisions). The left panel limits the analysis to subjects who live in areas where more than 29.32% of Hispanic individuals reside (70% of the sample). The right panel limits the analysis to subjects who live in areas where less than 29.32% of Hispanic individuals reside (30% of the sample). *Prior* is the belief of the employer about the minority group before the hiring set. *Experience* is the average productivity of the hired minority worker in each set. The estimations on Set 1 pool treatments HH and HB for posteriors about Hispanic workers and treatments BB and BH for posteriors about Black workers. Estimations on Set 2 use only the baseline treatments' data. This analysis restricts the sample to employers who hired at least one minority worker in both sets of hiring decisions, corresponding to 727 employers in treatments HH and HB (90.8% of the sample) and 731 employers in treatments BB and BH (91.4% of the sample). Robust standard errors are in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Appendix C Experiment materials

C.1 Study Advertisement

Figure C.1: Study description on Prolific



C.2 Employer Screenshots

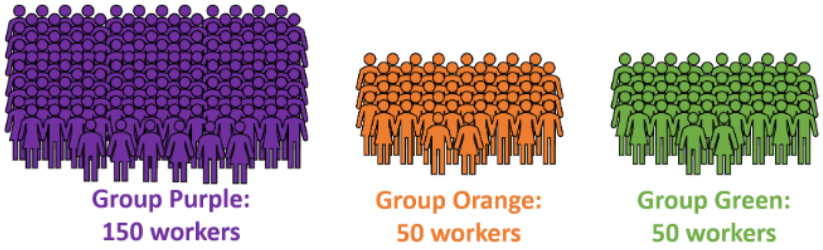
Figure C.2: General instructions

Your Hiring Decisions

This Section has two parts: Part 1 and Part 2. In both parts, you will act as an **employer** and you will have to make **hiring decisions** that consists in choosing which of two previous Prolific participants you want to hire.

The previous participants formed three different groups, based on their demographics.

- **White/Caucasian Americans** belong to **Group Purple**.
- **Hispanic/Latinos** belong to **Group Orange**.
- **Black/African Americans** belong to **Group Green**.



In Part 1, you will have to make **10 hiring decisions**. You will have to choose from which group you want to hire a worker. These previous participants in each group had to solve up to 10 multiple-choice algebraic math questions in 10 minutes. **You will earn money for every question the hired worker answered correctly**. In particular, once you choose which group you want to hire from**, you will earn 100 tokens for each problem solved correctly by the hired worker.**

IMPORTANT INFORMATION

All previous participants had to solve the same **10 multiple-choice algebraic math questions** in 10 minutes. On the next page, you will see 3 of these questions, and will have a chance to solve them for bonus earnings.

- The average number of problems solved correctly by **all the previous participants is 5 problems**.
- The average number of problems solved correctly by the workers from each group is **unknown to you**.

Every time you hire a worker, your **bonus earnings will be determined as follows**:
When you hire from Group Purple, **Group Orange, or Group Green**, you will be matched with a randomly selected worker and will then find out how many problems he or she solved correctly (between 0 and 10) and will earn 100 x the number of correctly solved problems.

Figure C.3: Belief elicitation example

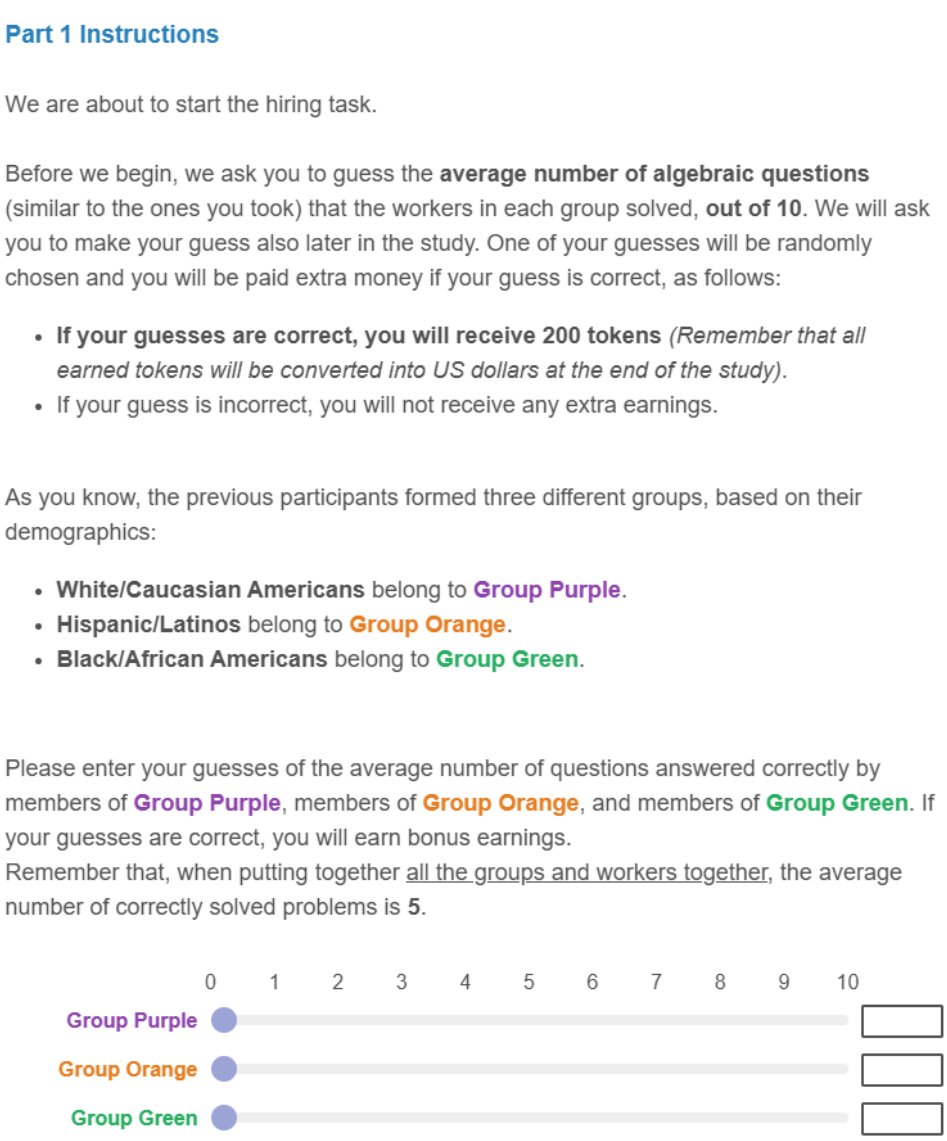


Figure C.4: Hiring decision - Set 1

Hiring Task - Part 1. Period 1

In this Part, you will make **10 hiring decisions**. Your earnings will be determined by 1 randomly selected hiring decision. Note that the hired and not hired workers will not be informed about your hiring decisions.

- **White/Caucasian Americans** belong to **Group Purple**.
- **Hispanic/Latinos** belong to **Group Orange**.
- **Black/African Americans** belong to **Group Green**.

In this Part of the study, you can only hire from **Group Purple** and from **Group Orange**.

Group Purple: 150 workers

Group Orange: 50 workers

Group Green: 50 workers

- If you hire from **Group Purple**, you will be matched with a randomly chosen worker from that group.
- If you hire from **Group Orange**, you will be matched with a randomly chosen worker from that group.
- You will be told the actual performance of the hired worker on the next page, earning you **100 x** the number of questions that the hired worker answered correctly, out of 10.
- Earnings are converted at the end of the study at the rate of **1,000 tokens for 1 USD.**

Please select the worker you want to hire:

Worker from Group Purple

Worker from Group Orange

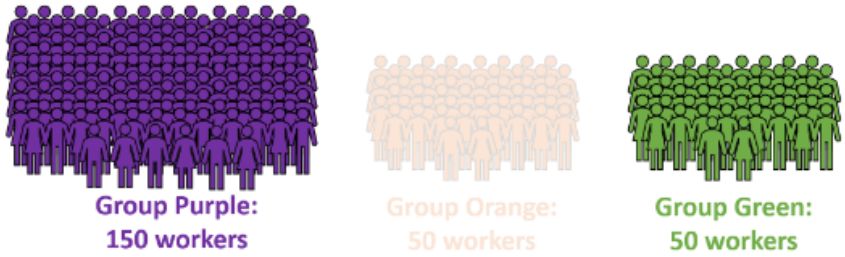
Figure C.5: Hiring decision - Set 2 (Spillover treatments)

Hiring Task - Part 2. Period 1

In this Part, you will make **10 hiring decisions**. Your earnings will be determined by 1 randomly selected hiring decision. Note that the hired and not hired workers will not be informed about your hiring decisions.

- **White/Caucasian Americans** belong to **Group Purple**.
- **Hispanic/Latinos** belong to **Group Orange**.
- **Black/African Americans** belong to **Group Green**.

In this Part of the study you can only hire from **Group Purple** and from **Group Green**.



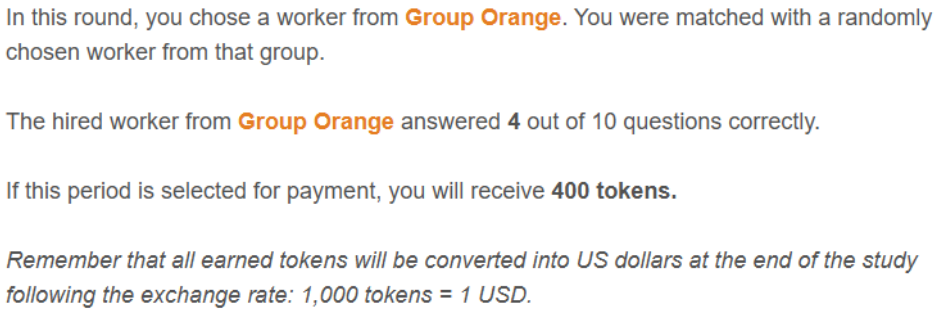
- If you hire from **Group Purple**, you will be matched with a randomly chosen worker from that group.
- If you hire from **Group Green**, you will be matched with a randomly chosen worker from that group.
- You will be told the actual performance of the hired worker on the next page, earning you **100 x** the number of questions that the hired worker answered correctly, out of 10.
- Earnings are converted at the end of the study at the rate of **1,000 tokens for 1 USD.**

Please select the worker you want to hire:


Worker from
Group Purple


Worker from
Group Green

Figure C.6: Feedback screen



C.3 Other survey questions

C.3.1 Measuring Social Desirability Bias

- Do you agree or disagree with the following statement? I'm always willing to admit it when I make a mistake.
  - Agree
  - Disagree
- Do you agree or disagree with the following statement? I like to gossip at times.
  - Agree
  - Disagree
- Do you agree or disagree with the following statement? There have been occasions when I took advantage of someone.
  - Agree
  - Disagree
- Do you agree or disagree with the following statement? I sometimes try to get even rather than forgive and forget.
  - Agree
  - Disagree
- Do you agree or disagree with the following statement? At times, I have really insisted on having things my own way.
  - Agree
  - Disagree

- Do you agree or disagree with the following statement? I have never been irked when people expressed ideas very different from my own.
  - Agree
  - Disagree
- Do you agree or disagree with the following statement? I have never deliberately said something that hurt someone's feelings.
  - Agree
  - Disagree

C.3.2 Example of Attention Checks

- What color is grass?
  The fresh, uncut grass, not leaves or hay. Make sure to select purple as an answer so that we know you are paying attention.
  - Green
  - Purple
- The color question you are about to respond to is very simple. When asked for your favorite color, you must select 'Green.' This is an attention check.
  Based on the text you read above, what color have you been asked to enter?
  - Red
  - Blue
  - Green
  - Orange
  - Brown
- Please select 'strongly disagree' in the statement below:
  I swim across the Atlantic Ocean to get to work every day.
  - Strongly Agree
  - Agree
  - Disagree
  - Strongly Disagree
  - Other

C.4 Ambiguity and Risk-aversion tasks

Figure C.7: Ambiguity aversion task

Part 4

For Part 4, you will imagine the following hypothetical situation:

- Consider a bucket containing **90 ping pong balls.**
- **30 balls** are **blue**.
- The remaining **60 balls** are either **red** or **yellow** in unknown proportions: it is not possible to know how many out of the 60 balls are red and how many are yellow.
- The balls are well mixed so that each individual ball is as likely to be drawn as any other.

We will present you two different scenarios. For each scenario, you will have to choose which option you prefer.

At the end of the study, one ball will be drawn from the bucket, and one of the scenarios will be randomly chosen. Your can **earn extra bonus earnings** depending on the randomly chosen scenario, the option that you choose in such scenario, and the color of the drawn ball.

Scenario 1.
Choose the option you prefer:

Option A. Receive 100 tokens if a **blue** ball is drawn, and 0 tokens otherwise.

Option B. Receive 100 tokens if a **red** ball is drawn, and 0 tokens otherwise.

Scenario 2.
Choose the option you prefer:

Option C. Receive 100 tokens if a **blue** ball or a **yellow** ball is drawn, and 0 tokens otherwise.

Option D. Receive 100 tokens if a **red** ball or a **yellow** ball is drawn, and 0 tokens otherwise.

At the end of the study, one ball will be randomly drawn from the bucket, and one of the two scenarios will be selected.
Your earnings will depend on the color of the ball, the selected scenario, and your chosen option.

Figure C.8: Risk aversion task

Part 5

For Part 5, you will **choose** one of six different gambles the gamble you would like to play. The six different gambles are listed in the table below.

- You must select one and only one of these gambles.
- To select a gamble, select the appropriate box.

Each gamble has two possible outcomes (ROLL LOW or ROLL HIGH) with the indicated probabilities of occurring. **Your earnings for this part** of the study will be determined by:

- Which of the six gambles you select; and
- Which of the two possible payoffs occur.

For example, if you select Gamble 4 and ROLL HIGH occurs, you will be paid 520 tokens. If ROLL LOW occurs, you will be paid 160 tokens. For every gamble, each ROLL has a 50% chance of occurring.

At the end of the study, a ten-sided die will be rolled to determine which event will occur. If the roll is 1, 2, 3, 4, or 5, ROLL LOW will occur. If the roll is 6, 7, 8, 9, or 0, ROLL HIGH will occur.

	Roll	Earnings	Chances
Gamble 1	LOW	280 tokens	50%
	HIGH	280 tokens	50%
Gamble 2	LOW	240 tokens	50%
	HIGH	360 tokens	50%
Gamble 3	LOW	200 tokens	50%
	HIGH	440 tokens	50%
Gamble 4	LOW	160 tokens	50%
	HIGH	520 tokens	50%
Gamble 5	LOW	120 tokens	50%
	HIGH	600 tokens	50%
Gamble 6	LOW	20 tokens	50%
	HIGH	700 tokens	50%

Appendix D Deviations from the Pre-registration

I pre-registered the study in AsPredicted, with reference number #107386 in September 2022. This Appendix outlines how the experimental design and data analysis in the paper deviate from the pre-registration.

Experimental Design

The pre-registration included a temporary policy that forced employers to hire minority workers. In particular, I specified treatment variations in which employers could only hire workers from the minority group for the first set of hiring decisions. The second set introduced freedom in hiring either with workers from the same group (and White workers), or with workers from a different minority group. While the pre-registration included a treatment with freedom of hiring in both sets, featuring a different minority group in the second set (ultimately becoming the Spillovers treatments), I deviated by not including the treatments with forced hiring, but instead featuring a treatment with freedom in hiring in both sets of hiring decisions, keeping the minority group available for employers constant (the Baseline treatments). This change corresponds to the need for a baseline treatment, where employers always face the same group of minority workers with freedom in hiring. Although not included in this paper, I studied the effects of a temporary hiring policy that forces employers to only hire minority workers, together with other hiring policies, as part of a different paper, Gomez-Vasquez (2025), pre-registered on AsPredicted with reference number #123012.

In the pre-registered design, I also stated that employers would make their hiring decisions between a White worker of average productivity and a randomly chosen minority worker. I deviated from the pre-registration by having employers' hiring decisions between a randomly chosen White worker and a randomly chosen minority Worker. This change corresponds to a more realistic approach, where employers do not know the productivity of the hired worker before hiring, even if the workers belong to the same racial or ethnic group as the employers. Nevertheless, restricting the availability of White workers in the labor market was studied as part of a different paper, Gomez-Vasquez et al. (2025), pre-registered on AsPredicted with reference number #184823.

Finally, the pre-registration only referred to treatments with racial and ethnic groups available for hire. The treatments featuring neutral groups, presented in Section 5.4.1, were not included. I expanded the experiment by adding the placebo experiment as a robustness check, keeping the same groups of workers from the main experiment, but without mentioning any references to racial or ethnic identities. Instead, employers are told that workers were assigned to groups Purple, Orange, and Green, without further explanations. I ran two treatments

for the placebo experiment: a baseline treatment in which employers can always hire between workers from Group Purple and workers from Group Orange, and a spillovers treatment in which workers from Group Green (instead of workers from Group Orange) were available in the second set of hiring decisions.

Estimation Strategy

While my estimation strategy followed the pre-registration by conducting regression analyses on the likelihood of hiring minority workers, as well as beliefs about the average productivity of minority workers, the pre-specified analysis included biases in beliefs as an outcome variable, and stated treatment dummies, dummies indicating positive or negative experiences with minority workers, and interaction terms between the treatment dummies and the positive/negative experience indicators, as primary explanatory variables.

I deviated from the pre-specified analysis in the analysis of beliefs by following a Bayesian learning framework, decomposing posterior beliefs about worker productivity into prior beliefs and hiring experiences. I conducted these estimations using the data from the baseline treatments and spillover treatments independently for each minority group. This was done to accurately estimate the role of priors and experiences with workers from one group on beliefs about the productivity of workers from the same group (baseline), and on beliefs about the productivity of workers from a different group (spillovers). Since there were two rounds of belief updating, i.e., employers state posterior beliefs after each set of hiring decisions, I conducted these estimations independently for each round of belief updating. Moreover, these estimations do not include an indicator variable for positive or negative hiring experiences. Instead, I included an extended version of the Bayesian learning framework for asymmetric belief updating featuring the interaction between the hiring surprise (the difference between hiring experience and prior beliefs) and an indicator for when the hiring surprise is negative. This interaction term represents the differential effect of a negative hiring experience compared to a positive hiring experience, captured by the uninteracted term. While these estimations do not include a treatment dummy, I included a pooled estimation for the baseline treatments with a dummy variable for the Hispanic group, equivalent to the treatment dummy (Table B.4).

My pre-specified analysis included the likelihood that a minority worker is hired in the first and second sets as a primary outcome. I extended this analysis by estimating the hiring of minority workers following a differences-in-differences approach, including an indicator for the second set of hiring decisions, a spillovers treatment indicator, and the interaction between the second set dummy and the treatment dummy. I also augmented my pre-specified analysis with the analysis of hiring dynamics, i.e., the role of hiring experiences on subsequent hiring. I followed a panel data approach with time fixed effects, estimating the impact of the experi-

ence with the currently hired minority worker on the number of minority workers hired in the subsequent periods using the actual performance of the hired worker, and a variation including indicator variables for positive and negative experiences. While this analysis was conducted for each minority group independently, I included a pooled estimation with a dummy variable for the Hispanic group, equivalent to the treatment dummy (Table B.5).

In addition, I augmented my pre-registered analysis in several ways. First, I conducted several robustness checks, contained in Section 5.4. Besides running the placebo experiment and conducting regression analyses including interaction terms of the explanatory variables and an indicator for the placebo experiments, I addressed concerns on experimenter demand effect, inattention, and social desirability bias. For experimenter demand effects, I identified the subjects that correctly guessed the purpose of the experiment (according to an open-ended question at the end of the experiment), and test whether the results hold excluding them. I conduct a similar procedure for inattention, and test whether the results hold after excluding subjects who failed one attention check (subjects who failed two or more were excluded from the study, as stated in the pre-registration). While I pre-registered social desirability bias as a control, I tested for heterogeneous effects based on the social desirability score following Dhar et al. (2022). Second, while I deviated from the pre-registration by not including biases in beliefs as an outcome variable, defined as the difference between beliefs about minority workers' performance and White workers' performance, I included the analysis on beliefs about White workers' performance. I also augmented the analysis by including the hiring dynamics of White workers, replicating the analysis done for minority workers. Third, I included heterogeneity analysis based on initial bias against minority workers. Specifically, I replicated the main analysis for two subsamples based on "biased" (those who had lower initial priors about the performance of both Black and Hispanic workers than about White workers) and "unbiased" employers.

Finally, I added an exploratory analysis of possible mechanisms behind the results. On the one hand, I decomposed the sources of discrimination in an attempt to disentangle taste-based and (inaccurate) statistical discrimination, following Bohren et al. (2025). On the other hand, I assess whether the proportion of Black and Hispanic individuals in the location of study participants affects their initial priors and probability of hiring of minority workers, and then I employ threshold regressions to investigate whether such proportions affect the belief updating process. While I only report the results for the proportion of Hispanic population and beliefs about Hispanic workers, I also conducted the analysis for beliefs about Black workers and the proportion of Black population. These results are noted in footnote 47. These additional analyses of mechanisms are not pre-registered but are crucial for studying the main mechanisms expected to drive the results, given the naturally occurring racial and ethnic identities studied in the experiment.