# Misperception and informativeness in statistical discrimination*

Matteo Escudé
LUISS

Paula Onuchic
London School of Economics

Ludvig Sinander
University of Oxford

Quitzé Valenzuela-Stookey
UC Berkeley

27 August 2025

**Abstract**

We study the interplay of information and prior (mis)perceptions in a Phelps–Aigner–Cain-type model of statistical discrimination in the labor market. We decompose the effect on average pay of an increase in how informative observables are about workers' skill into a non-negative *instrumental* component, reflecting increased surplus due to better matching of workers with tasks, and a *perception-correcting* component capturing how extra information diminishes the importance of prior misperceptions about the distribution of skills in the worker population. We sign the perception-correcting term: it is non-negative (non-positive) if the population was ex-ante under-perceived (over-perceived). We then consider the implications for pay gaps between equally-skilled populations that differ in information, perceptions, or both, and identify conditions under which improving information narrows pay gaps.

## 1 Introduction

There are significant pay gaps between populations, for example between women and men and between different ethnic groups.[1] Possible contributing

*We are grateful for useful comments from Aislinn Bohren, Davide Bordoli, Chris Chambers, Federico Echenique, Ben Golub, Ian Jewitt, Meg Meyer, Augustus Smith, and the audience at the LSE Perspectives on Economic Theory conference. Oscar Calvert provided excellent research assistance.

[1]See e.g. U.S. Department of Labor (2025) for the United States and Office for National Statistics (2023, 2024a, 2024b) for the United Kingdom.

factors include discrimination, differences in human capital (potentially caused by discrimination earlier in life), and differences in individual behavior, such as occupation choice (also potentially influenced by discrimination).

The literature on *statistical discrimination* argues that these pay gaps are partly explained by information-economic forces.[2] Statistical discrimination is possible whenever workers' skill is imperfectly observable. One way in which this can produce pay gaps is through *misperception,* whereby firms believe that one population is less skilled on average than is actually the case. For example, users of the website StackExchange appear to harbor misperceptions about women's and men's abilities (Bohren, Imas & Rosenberg, 2019).[3] Another way in which skill unobservability can generate pay gaps is through differential *informativeness,* when workers' observable characteristics (e.g. CVs and test scores) are less predictive of skill in one population than in another. For example, SAT scores are less predictive of academic ability for poorer students than for richer ones (Rothstein, 2004).

In this paper, we study the interplay between misperception and informativeness in determining pay gaps, using a simple but general labor-market model in the spirit of the statistical-discrimination literature following Phelps (1972a, 1972b) and Aigner and Cain (1977).[4] The most-emphasized result in the literature (see Aigner & Cain, 1977) is that a population's average pay is higher if its observable characteristics are more informative about skill. This finding relies on the assumption that perceptions are accurate. As for varying the perceptions themselves, it is intuitive and in fact true that a population is paid more on average the more favorably it is perceived.

Our starting point is the observation that these two effects interact: in particular, a population that is *both* more informative *and* more favorably perceived may be paid *strictly less* on average. To see why, consider an economy in which every worker is either a "high type" with productivity 1 or a "low type" with productivity 0. Let $p \in (0, 1)$ and $q \in (0, 1)$ denote, respectively, the true and perceived fractions of high types. Assume that labor demand is competitive, so that each worker's pay equals her posterior expected productivity. If observable characteristics are completely uninform-

[2]See e.g. Lang and Lehmann (2012), Azmat and Petrongolo (2016), Bertrand and Duflo (2017), Blau and Kahn (2017), Neumark (2018) and Lang and Spitzer (2020). A related but distinct literature on "taste-based" discrimination (Becker, 1955, 1957) seeks explanations rooted in decision-makers' biased preferences, rather than in information.

[3]Similarly, Agan and Starr (2018) and Arnold, Dobbie and Yang (2018) present evidence of misperceptions (held by employers and by bail judges, respectively) about racial differences in the rates of commission of and conviction for crime in the United States.

[4]Specifically, we adapt the model of Chambers and Echenique (2021).

ative about productivity, then all workers are paid $q$. If observables perfectly reveal productivity, then high types are paid 1 and low types are paid 0, so average pay is $p \cdot 1 + (1-p) \cdot 0 = p$. Thus if the population is over-perceived, i.e. $q > p$, then average pay is strictly *lower* if observables perfectly reveal skill than if they are completely uninformative. The reason why this happens is that information corrects misperceptions, which is harmful for an over-perceived population. The implication for pay gaps is that between a perfectly informative population $I$ and a completely uninformative population $J$, the pay gap $p - q_J$ is negative whenever population $J$ is over-perceived ($q_J > p$), even if population $I$ is more favorably perceived than population $J$ ($q_I > q_J$).

Our main theoretical contribution is to unpack this interaction between information and misperception. In particular, we decompose the change in a population's average pay when observables become more informative about skill into two terms. The first term, which we call *instrumental,* captures how extra information is used to improve the assignment of workers to tasks, thereby increasing expected surplus and (thus) pay.[5] The second term, the *perception-correcting component,* captures how the increased availability of information about workers diminishes the influence of prior misperceptions about the distribution of skills in the population. This term can be either negative or positive. (From the previous paragraph, we know that it can be negative enough to dominate the never-negative instrumental term.) Our theorem signs the perception-correcting component, showing that it is non-negative if the population was ex-ante under-perceived, non-positive if it was over-perceived, and zero if it was accurately perceived.

We then apply our decomposition to statistical discrimination, meaning pay gaps between populations whose (true) skill distributions are equal. In particular, we identify conditions under which a population $I$ that is both more informative than and more favorably perceived than another population $J$ will be paid more on average. The key condition turns out to be whether population $J$ is under-perceived (relative to the true skill distribution): if yes, then population $I$ will indeed earn more on average, and if no, then the reverse may occur (as shown above by example).

The spirit of the perception-correcting term is that prior perceptions about a population matter less when more information is available about individual workers. This suggests that the pay gap between two populations will narrow when more information is made available about workers in both populations, for example due to disclosure mandates or technological

[5] This term is zero in the example in the previous paragraph, since that example involves no task assignment.

progress. Bohren, Imas and Rosenberg (2019) confirm this intuition in a model of subjective evaluation (rather than of pay) that can be viewed as a special case of our model. We assess the validity of the "information narrows pay gaps" intuition beyond this special case, identifying (stringent) conditions under which it is correct, and showing that extra information may *widen* pay gaps when these conditions do not hold. An implication is that informativeness-increasing policy interventions can backfire.

### 1.1 Related literature

We contribute to the statistical-discrimination literature initiated by Phelps (1972a, 1972b) and Aigner and Cain (1977), which shows how pay gaps can arise from firms' inference problem of estimating workers' unobservable skills from observables such as test scores and CVs.[6] Within this literature, we take inspiration from Chambers and Echenique (2021), borrowing (and slightly enriching) their simple but general "Phelpsian" model of the labor market.

Most work on statistical discrimination has focused on the impact on pay of differential informativeness of observables about skill, under the assumption that firms accurately perceive the distribution of skills.[7] More recent work has argued for the importance of misperceptions in explaining empirical patterns of discrimination (e.g. Bohren, Imas & Rosenberg, 2019; Bohren, Haggag, Imas & Pope, 2025).[8] Our analysis elucidates how informativeness and misperception interact to determine pay gaps.

The instrumental component in our decomposition captures the familiar instrumental value of information, defined and characterized by Blackwell (1951, 1953).[9] The idea of a perception-correcting appears in the statistics literature,[10] and plays a role in, for example, persuasion (e.g. Alonso & Câmara, 2016; Onuchic & Ray, 2023) and behavioral decision theory (Bordoli, 2025). Bordoli's paper in particular distinguishes (informally) between instrumental and non-instrumental effects of new information. Our results on signing the perception-correcting component are related to Kartik, Lee and Suen (2021), as we discuss in section 3.

[6] A related but distinct literature following Arrow (1973), also called "statistical discrimination," studies pay gaps arising from "bad equilibria" rather than from inference. For an overview of both theoretical literatures, see Fang and Moro (2011) and Onuchic (2023).

[7] See the systematic literature review by Bohren, Haggag, Imas and Pope (2025).

[8] One could also consider misperceptions about the structure of correlation between observables and skill. We discuss this possibility in section 6.

[9] Blackwell's theorem is stated in section 4.1 below.

[10] See Laplace (1774), Wald (1947), Girsanov (1960) and Blackwell and Dubins (1962).

### 1.2 Roadmap

We describe the model in section 2, and present our decomposition result in section 3. We then draw conclusions about when there are pay gaps between equally-skilled populations (section 4), and whether these narrow as more information is made available (section 5). All proofs are in the appendix.

## 2 Model

We consider a model in the spirit of Phelps (1972b) and Aigner and Cain (1977). Specifically, we adopt the model of Chambers and Echenique (2021), and enrich it slightly by allowing for prior misperceptions.[11]

The surplus generated by a worker at a firm depends on her abilities and on the task to which she is assigned. Formally, each worker has a *skill type* $\theta$ drawn from a finite set $\Theta \subset \mathbb{R}$ with $|\Theta| \geqslant 2$. A *task* is a vector $a \in \mathbb{R}^{\Theta}$, where the interpretation is that $a(\theta) \in \mathbb{R}$ is how much surplus is generated (in expectation) when a worker of skill type $\theta \in \Theta$ performs task $a$. The set of all tasks is denoted $\mathcal{A} \coloneqq \mathbb{R}^{\Theta}$.

Firms are described by their technology, meaning what tasks are available. Formally, a *firm* is a non-empty finite set $A \subset \mathcal{A}$. A firm $A$ is called *monotone* if $A \subset \mathcal{A}_M \coloneqq \{a \in \mathbb{R}^{\Theta} : a(\theta') > a(\theta) \text{ whenever } \theta' > \theta\}$, meaning that each task's surplus is increasing in the worker's skill. Considering all firms $A \subset \mathcal{A}$ amounts to assuming that skill types are horizontally differentiated, whereas considering only monotone firms $A \subset \mathcal{A}_M$ means that skill types are vertically differentiated (the higher a worker's skill type, the more productive she is at *every* task). We primarily focus on monotone firms.

Worker skill is unobservable, so firms must estimate skill based on observables. We describe a worker's observables by a *signal* $s$. The signal should be thought of as a vector of observable characteristics, perhaps including a CV and standardized test scores. Formally, the worker population's *signal structure* is a pair $\langle S, \pi \rangle$, where $S$ is a non-empty finite set and $\pi : S \times \Theta \to [0, 1]$ satisfies $\sum_{s \in S} \pi(s|\theta) = 1$ for each skill type $\theta \in \Theta$.[12] The interpretation is that $S$ is the set of possible signals $s$ (e.g. possible combinations of CV contents and test scores), and that $\pi(s|\theta)$ is the probability that a type-$\theta$ worker would have signal $s$. We assume that for every signal $s \in S$, there is at least one skill type $\theta \in \Theta$ such that $\pi(s|\theta) > 0$.[13]

[11] Chambers and Echenique's results are briefly discussed in section 4.1 below.

[12] Signal structures are also called "information structures" or "(Blackwell) experiments."

[13] This assumption is without loss of generality, because if there were a signal $s \in S$ with $\pi(s|\theta) = 0$ for every $\theta \in \Theta$, then we could neglect $s$ entirely, by deleting it from $S$.

Firms may misperceive the distribution of skill types in the population: firms' subjective prior *perception* $q \in \Delta(\Theta)$ of the skill type distribution may differ from the true distribution $p \in \Delta(\Theta)$. For simplicity, we assume that both $q$ and $p$ have full support.

To estimate a worker's ability based on her signal $s \in S$, firms apply Bayes's rule, informed by their prior perception $q \in \Delta(\Theta)$ and by the signal structure $\langle S, \pi \rangle$. Concretely, the posterior probability which firms assign to a worker with signal $s \in S$ having skill type $\theta \in \Theta$ is

$$q_{\langle S,\pi\rangle}(\theta|s) \coloneqq \frac{q(\theta)\pi(s|\theta)}{\sum_{\theta' \in \Theta} q(\theta')\pi(s|\theta')}.$$

Each firm $A \subset \mathcal{A}$ produces efficiently, allocating each worker to whichever task $a \in A$ yields the highest expected surplus given that worker's signal $s \in S$, where the expected-surplus calculations are based on the perception $q$ and the signal structure $\langle S, \pi \rangle$. Thus the maximized expected surplus calculated by a firm $A \subset \mathcal{A}$ faced with a worker with signal $s \in S$ is

$$w_A(s, q, \langle S, \pi \rangle) \coloneqq \max_{a \in A} \sum_{\theta \in \Theta} q_{\langle S,\pi\rangle}(\theta|s) a(\theta).$$

Labor demand is competitive, so each worker is paid the expected surplus that she generates.[14] Average pay in the population is therefore

$$W_A(p, q, \langle S, \pi \rangle) \coloneqq \sum_{\theta \in \Theta} p(\theta) \sum_{s \in S} \pi(s|\theta) w_A(s, q, \langle S, \pi \rangle).$$

Note the differing roles of $p$ and $q$: a worker's pay depends on firms' prior perception $q$ (via the posterior perception $q_{\langle S,\pi\rangle}$), but the averaging of different workers' pay is according to the true distribution $p$.

To analyze discrimination, we compare average pay across two separate populations $I$ and $J$ with respective (true) skill distributions $p_I$ and $p_J$, perceptions $q_I$ and $q_J$, and signal structures $\langle S_I, \pi_I \rangle$ and $\langle S_J, \pi_J \rangle$. Each worker's population membership ($I$ or $J$) is observable. Our focus is on statistical discrimination in the labor market: how the gap in average pay between the two populations is influenced by the perceptions $q_I, q_J$ and signal structures $\langle S_I, \pi_I \rangle, \langle S_J, \pi_J \rangle$. In order to isolate statistical discrimination in the labor market from the important but separate issue of human-capital inequalities arising earlier in life, we always consider populations with equal skill distributions: $p_I = p_J = p$.[15]

[14]Our conclusions do not change if workers are instead paid a fixed (i.e. signal-independent) fraction $\alpha \in (0, 1)$ of their expected surplus.

[15]For the same reason, we rule out taste-based discrimination; this is implicit in our assumption that workers (in any population) are paid their expected surplus.

We compare any two perceptions $q$ and $q'$ in terms of their *favorableness* in the likelihood-ratio order: $q' \succsim_{LR} q$ if $q(\theta)q'(\theta') \geqslant q(\theta')q'(\theta)$ holds whenever $\theta' > \theta$. We say that the population is *under-perceived* if $p \succsim_{LR} q$, and *over-perceived* if $q \succsim_{LR} p$.

We compare any two signal structures $\langle S, \pi \rangle$ and $\langle S', \pi' \rangle$ in terms of their *informativeness* in the garbling sense (Blackwell, 1951, 1953): $\langle S', \pi' \rangle$ is more informative than $\langle S, \pi \rangle$, written $\langle S', \pi' \rangle \succsim_G \langle S, \pi \rangle$, if there exists a garbling kernel from $\langle S', \pi' \rangle$ to $\langle S, \pi \rangle$, i.e. a map $g : S \times S' \to [0, 1]$ satisfying $\sum_{s \in S} g(s|s') = 1$ for each $s' \in S'$ such that $\pi(s|\theta) = \sum_{s' \in S} g(s|s')\pi'(s'|\theta)$ for each $s \in S$ and $\theta \in \Theta$. A signal structure $\langle S, \pi \rangle$ is said to be *MLR (monotone likelihood ratio)* if $S \subset \mathbb{R}$ and $\pi(s|\theta)\pi(s'|\theta') \geqslant \pi(s|\theta')\pi(s'|\theta)$ holds whenever $s' > s$ and $\theta' > \theta$.

**Remark 1.** An alternative interpretation of our model is that $q$ is the true distribution of skill in the population, which firms perceive correctly, and that $p$ is the distribution of skill in a sub-population of interest (or a re-weighting of the full population according to some welfare weights[16]). On this interpretation, when comparing two populations $I$ and $J$ with different skill distributions $q_I \neq q_J$, the comparison is made "like-for-like," between two sub-populations with equal skill distributions $p_I = p_J = p$.

## 2.1 A preliminary result

Intuition suggests that whatever the true skill distribution $p$ and the signal structure $\langle S, \pi \rangle$, a population's average pay will be higher the more favorably it is perceived ex ante. The following lemma formalizes this intuition, and furnishes a converse. We use this result throughout the paper.

**Lemma 1.** Fix a skill distribution $p$, and consider two perceptions $q$ and $q'$.

(a) If $q'$ is more favorable than $q$ ($q' \succsim_{LR} q$), then for any signal structure $\langle S, \pi \rangle$,

$$W_A(p, q', \langle S, \pi \rangle) \geqslant W_A(p, q, \langle S, \pi \rangle) \quad \text{for any monotone firm } A \subset \mathcal{A}_M.$$

(b) If $q'$ is not more favorable than $q$ ($q' \not\succsim_{LR} q$), then there exists a signal structure $\langle S, \pi \rangle$ such that

$$W_A(p, q', \langle S, \pi \rangle) < W_A(p, q, \langle S, \pi \rangle) \quad \text{for any monotone firm } A \subset \mathcal{A}_M.$$

[16] As in Bergson (1938) and Samuelson (1947), and more recently in e.g. Dworczak, Kominers and Akbarpour (2021).

## 3 Decomposing the impact of new information

In this section, we present our main theoretical contribution: a decomposition of the impact of new information on average pay into two components reflecting distinct *instrumental* and *perception-correcting* effects, and a result which signs these two components. The instrumental effect captures how firms use extra information to tailor task assignment more finely to workers' likely skills. The perception-correcting effect reflects the diminished importance of any misperception $q \neq p$ as information becomes more precise; indeed, in the extreme case in which the signal structure becomes fully informative, the perception $q$ ceases entirely to matter for pay, since each worker's skill type is revealed and she is paid accordingly.

To describe our decomposition, fix a firm $A \subset \mathcal{A}$, a skill distribution $p$, and a perception $q$, and consider a shift from signal structure $\langle S, \pi \rangle$ to $\langle S', \pi' \rangle$, where the latter is more informative (that is, $\langle S', \pi' \rangle \succsim_G \langle S, \pi \rangle$). Then there exists a joint distribution of signals and skill such that the distribution of $\theta$ conditional on both $s$ and $s'$ is the same as that conditional on $s'$ alone (in other words, $\theta$ is independent of $s$ conditional on $s'$). Specifically, that joint distribution is $\mu_p \in \Delta(\Theta \times S \times S')$ given by

$$\mu_p(\theta, s, s') := p(\theta)\pi'(s'|\theta)g(s|s') \quad \text{for all } \theta \in \Theta,\ s \in S \text{ and } s' \in S',$$

where $g$ is the garbling kernel from $\langle S', \pi' \rangle$ to $\langle S, \pi \rangle$. From the perspective of firms, who believe that skill is distributed according to $q$ rather than $p$, the corresponding joint distribution is $\mu_q \in \Delta(\Theta \times S \times S')$ given by

$$\mu_q(\theta, s, s') := q(\theta)\pi'(s'|\theta)g(s|s') \quad \text{for all } \theta \in \Theta,\ s \in S \text{ and } s' \in S'.$$

For each signal $s \in S$ of the less informative signal structure $\langle S, \pi \rangle$, let

$$\widehat{a}_s \in \arg\max_{a \in A} \sum_{\theta \in \Theta} q_{\langle S, \pi \rangle}(\theta|s) a(\theta)$$

denote firm $A$'s surplus-maximizing task choice for a worker with signal $s$. Similarly, for each $s' \in S'$, let

$$\widehat{a}'_{s'} \in \arg\max_{a \in A} \sum_{\theta \in \Theta} q_{\langle S', \pi' \rangle}(\theta|s') a(\theta)$$

denote a surplus-maximizing task choice under the more informative signal structure $\langle S', \pi' \rangle$.

We shall decompose the change $W_A(p, q, \langle S', \pi' \rangle) - W_A(p, q, \langle S, \pi \rangle)$ in average pay into two terms. The first term, which we call *perception-correcting,* captures the effect on average pay of the change in beliefs induced by the extra information contained in $\langle S', \pi' \rangle$, *holding fixed* the firm's task-assignment choices $s \mapsto \widehat{a}_s$. The second term, called *instrumental,* captures the increase of expected surplus from firms using the extra information in $\langle S', \pi' \rangle$ to better tailor task assignment to workers' likely skills, *holding perceptions fixed.* The formal definitions of these two effects are as follows.

**Definition 1.** For any firm $A \subset \mathcal{A}$, skill distribution $p$, perception $q$, and signal structures $\langle S', \pi' \rangle \succsim_G \langle S, \pi \rangle$, the *perception-correcting component* of the change in average pay is

$$\begin{aligned} &\mathcal{C}_A(p, q, \langle S, \pi \rangle, \langle S', \pi' \rangle) \\ &\qquad \coloneqq \sum_{s \in S} \mu_p(s) \sum_{s' \in S'} \left[\mu_p(s'|s) - \mu_q(s'|s)\right] \sum_{\theta \in \Theta} \mu_q(\theta|s, s') \widehat{a}_s(\theta), \end{aligned}$$

and the *instrumental component* of the change in average pay is

$$\begin{aligned} &\mathcal{I}_A(p, q, \langle S, \pi \rangle, \langle S', \pi' \rangle) \\ &\qquad \coloneqq \sum_{s' \in S'} \mu_p(s') \sum_{\theta \in \Theta} q_{\langle S', \pi' \rangle}(\theta|s') \left[\widehat{a}'_{s'}(\theta) - \sum_{s \in S} g(s|s') \widehat{a}_s(\theta)\right] \\ &\qquad = \sum_{s \in S} \mu_p(s) \sum_{s' \in S'} \mu_p(s'|s) \sum_{\theta \in \Theta} \mu_q(\theta|s, s') \left[\widehat{a}'_{s'}(\theta) - \widehat{a}_s(\theta)\right]. \end{aligned}$$

Our main result asserts that the change in average pay decomposes into perception-correcting and instrumental components, and that these components may be signed.

**Theorem 1.** Fix a firm $A \subset \mathcal{A}$, a skill distribution $p$, a perception $q$, and signal structures $\langle S', \pi' \rangle \succsim_G \langle S, \pi \rangle$.

(a) The change in average pay admits the decomposition

$$\begin{aligned} &W_A(p, q, \langle S', \pi' \rangle) - W_A(p, q, \langle S, \pi \rangle) \\ &\qquad = \mathcal{C}_A(p, q, \langle S, \pi \rangle, \langle S', \pi' \rangle) + \mathcal{I}_A(p, q, \langle S, \pi \rangle, \langle S', \pi' \rangle). \end{aligned}$$

(b) The instrumental component $\mathcal{I}_A(p, q, \langle S, \pi \rangle, \langle S', \pi' \rangle)$ is non-negative.

Suppose in addition that the firm is monotone ($A \subset \mathcal{A}_M$) and that the more informative signal structure $\langle S', \pi' \rangle$ is MLR.

(c) If the population is under-perceived ($p \succsim_{LR} q$), then the perception-correcting component $\mathcal{C}_A(p, q, \langle S, \pi \rangle, \langle S', \pi' \rangle)$ is non-negative.

(d) If the population is over-perceived ($q \gtrsim_{LR} p$), then the perception-correcting component $\mathcal{C}_A(p, q, \langle S, \pi \rangle, \langle S', \pi' \rangle)$ is non-positive.

The non-negativity of the instrumental component (part (b)) is a slight generalization of the "easy" half of Blackwell's (1951, 1953) theorem, which asserts the same conclusion under the additional hypothesis that the perception $q$ is accurate, i.e. equal to the true distribution $p$. The familiar intuition is that extra information cannot hurt a decision-maker, since she can always ignore it. (Blackwell's theorem is stated in full in section 4.1 below.)

This familiar intuition is invalid in our model, because extra information also has the secondary effect of partially correcting any prior misperception $q \neq p$. The definition of the instrumental component $\mathcal{I}_A(p, q, \langle S, \pi \rangle, \langle S', \pi' \rangle)$ shuts down this effect by holding perceptions fixed, thereby restoring the usual non-negativity of the instrumental value of information (part (b)).

The secondary effect is isolated in the perception-correcting component, in which task assignment is held fixed. This term is zero if the prior perception $q$ is accurate, i.e. equal to the true distribution. For an under-perceived population, the perception-correcting effect is non-negative (part (c)) because as more information becomes available, firms recognize that the skill distribution is better than their prior perception $q$ suggested, leading them to pay workers more. The same logic explains why the perception-correcting term is non-positive for an over-perceived population (part (d)). These intuitions are incomplete because they do not exploit the theorem's monotonicity and MLR assumptions, without which the result may fail, as we show in section 3.2 below.

More abstractly, the force behind the perception-correcting term is that posterior beliefs about workers' skill are (bary)centered somewhere between $q$ and $p$, and that this point is further from $q$ and closer to $p$ the more informative is the signal structure (it is $q$ under no information, and $p$ under full information). This reflects the fact that Bayesians are not dogmatic, but rather update beliefs about a population in light of the evidence.

The idea of a perception-correcting effect has appeared in various contexts, such as statistics,[17] persuasion with heterogeneous priors (e.g. Alonso & Câmara, 2016; Onuchic & Ray, 2023), and behavioral decision theory (Bordoli, 2025). Bordoli's paper in particular distinguishes between "instrumental" and "non-instrumental" effects of information. Closest to Theorem 1(c)–(d) is the result of Kartik, Lee and Suen (2021), which formalizes a sense in which posterior beliefs become (bary)centred further from $q$ and closer to $p$

[17] See Laplace (1774), Wald (1947), Girsanov (1960) and Blackwell and Dubins (1962).

as the signal structure becomes more informative. That result is formally a special case of Theorem 1(c)–(d), and the proofs are similar.

### 3.1 Signing the total impact of new information

Theorem 1 can sometimes be used to determine the sign of the total impact $W_A(p, q, \langle S', \pi' \rangle) - W_A(p, q, \langle S, \pi \rangle)$ of new information on average pay.

**Corollary 1.** Fix a skill distribution $p$, a perception $q$, and signal structures $\langle S', \pi' \rangle \gtrsim_G \langle S, \pi \rangle$, where $\langle S', \pi' \rangle$ is MLR. If the population is under-perceived ($p \gtrsim_{LR} q$), then new information increases average pay:

$$W_A(p, q, \langle S', \pi' \rangle) \geqslant W_A(p, q, \langle S, \pi \rangle) \quad \text{for any monotone firm } A \subset \mathcal{A}_M.$$

Conversely, if the population is over-perceived ($q \gtrsim_{LR} p$), so that the instrumental and perception-correcting components have opposite signs, then the total effect is ambiguous and can be negative, as the following example shows. (This example was discussed in the introduction.)

**Example 1.** Suppose that skill types are binary, $\Theta = \{0, 1\}$. Let $\langle S, \pi \rangle$ be an uninformative signal structure: $\pi(s|0) = \pi(s|1)$ for every $s \in S$. Let $\langle S', \pi' \rangle$ be fully informative: $\pi'(s^0 \mid 0) = \pi'(s^1 \mid 1) = 1$ for some $s^0 \neq s^1$ in $S'$. Obviously these signal structures are MLR and satisfy $\langle S', \pi' \rangle \succ_G \langle S, \pi \rangle$.

Let $\overline{a} \in \mathcal{A}_M$ be the task given by $\overline{a}(\theta) := \theta$ for each $\theta \in \Theta$, and consider the monotone firm $A = \{\overline{a}\}$. This firm pays each worker their posterior probability of being the high skill type: $w_{\{\overline{a}\}}(s'', q, \langle S'', \pi'' \rangle) = q_{\langle S'', \pi'' \rangle}(1|s'')$ for any signal structure $\langle S'', \pi'' \rangle$ and signal $s'' \in S''$. Hence the change in average pay is

$$\begin{aligned} &W_{\{\overline{a}\}}(p, q, \langle S', \pi' \rangle) - W_{\{\overline{a}\}}(p, q, \langle S, \pi \rangle) \\ &\qquad = \left[p(0) q_{\langle S', \pi' \rangle}\left(1 \middle| s^0\right) + p(1) q_{\langle S', \pi' \rangle}\left(1 \middle| s^1\right)\right] - q(1) = p(1) - q(1). \end{aligned}$$

Thus if the population is strictly over-perceived ($q(1) > p(1)$), then adding information strictly *decreases* average pay.

By Theorem 1, such a "reversal" can occur only if the perception-correcting component is sufficiently negative to dominate the always non-negative instrumental component. To verify this directly, observe that the instrumental component is zero, since that term captures the use of new information to change task assignment, and there is only one task available:

$$\begin{aligned} &\mathcal{I}_{\{\overline{a}\}}(p, q, \langle S, \pi \rangle, \langle S', \pi' \rangle) \\ &\qquad = \sum_{s \in S} \mu_p(s) \sum_{s' \in S'} \mu_p(s'|s) \sum_{\theta \in \Theta} \mu_q(\theta|s, s')[\overline{a}(\theta) - \overline{a}(\theta)] = 0. \end{aligned}$$

The perception-correcting term, on the other hand, is equal to

$$\begin{aligned}
&\mathcal{C}_{\{\overline{a}\}}(p, q, \langle S, \pi\rangle, \langle S', \pi'\rangle) \\
&\qquad = \sum_{s\in S} \mu_p(s) \sum_{s'\in S'} [\mu_p(s'|s) - \mu_q(s'|s)] \sum_{\theta\in\Theta} \mu_q(\theta|s, s')\overline{a}(\theta) \\
&\qquad = \sum_{\theta'\in\Theta} \left[\mu_p\left(s^{\theta'}\right) - \mu_q\left(s^{\theta'}\right)\right] \sum_{\theta\in\Theta} \mu_q\left(\theta\middle|s^{\theta'}\right)\overline{a}(\theta) \\
&\qquad = \sum_{\theta'\in\Theta} [p(\theta') - q(\theta')]\overline{a}(\theta') = p(1) - q(1).
\end{aligned}$$

## 3.2 The role of the monotonicity and MLR assumptions

In Theorem 1(c)–(d), it is assumed that the firm under consideration is monotone and that the more informative of the two signal structures is MLR. The following two examples illustrate the role of these assumptions by showing how without them, the conclusions of Theorem 1(c)–(d) may fail.

**Example 2.** Consider Example 1, except with a different firm: let $\underline{a} \in \mathcal{A} \setminus \mathcal{A}_M$ be the task given by $\underline{a}(\theta) \coloneqq 1 - \theta$ for each $\theta \in \Theta$, and consider the non-monotone firm $A = \{\underline{a}\}$.[18] This firm pays each worker their posterior probability of being the low skill type: $w_{\{\underline{a}\}}(s'', q, \langle S'', \pi''\rangle) = q_{\langle S'',\pi''\rangle}(0|s'')$ for any signal structure $\langle S'', \pi''\rangle$ and signal $s'' \in S''$. Hence the change in average pay is

$$\begin{aligned}
&W_{\{\underline{a}\}}(p, q, \langle S', \pi'\rangle) - W_{\{\underline{a}\}}(p, q, \langle S, \pi\rangle) \\
&\qquad = \left[p(0)q_{\langle S',\pi'\rangle}\left(0\middle|s^0\right) + p(1)q_{\langle S',\pi'\rangle}\left(0\middle|s^1\right)\right] - q(0) = p(0) - q(0).
\end{aligned}$$

The instrumental component $\mathcal{I}_A(p, q, \langle S, \pi\rangle, \langle S', \pi'\rangle)$ is zero since task assignment does not change (as firm $\{\underline{a}\}$ has only one task), so the perception-correcting component is $\mathcal{C}_A(p, q, \langle S, \pi\rangle, \langle S', \pi'\rangle) = p(0)-q(0)$ by Theorem 1(a). Hence the perception-correcting component is strictly negative if the population is under-perceived ($q(0) > p(0)$) and strictly positive if the population is over-perceived ($q(0) < p(0)$). This shows that the monotonicity hypothesis cannot be dispensed with in Theorem 1(c)–(d).

**Example 3.** Suppose that skill types are ternary, $\Theta = \{0, 1, 2\}$, and consider the perception $q$ given by $q(0) = q(1) = 1/4$ and $q(2) = 1/2$. Let $\langle S, \pi\rangle$ be an uninformative signal structure: $\pi(s|0) = \pi(s|1) = \pi(s|2)$ for every $s \in S$. Let

[18] The firm $\{\underline{a}\}$ is not monotone since by definition, a "monotone" firm $A$ is one whose every task $a \in A$ is strictly *increasing*.

$\langle S', \pi' \rangle$ reveal whether skill is 1 and nothing else: $\pi'(s^1 \mid 1) = \pi'(s^{0,2} \mid 0) = \pi'(s^{0,2} \mid 2) = 1$ for some $s^1 \neq s^{0,2}$ in $S'$. By inspection, $\langle S', \pi' \rangle \succ_G \langle S, \pi \rangle$, and the signal structure $\langle S', \pi' \rangle$ is not MLR. Posterior beliefs about skill are $q_{\langle S,\pi \rangle}(\cdot|s) = q(\cdot)$ for every $s \in S$, $q_{\langle S',\pi' \rangle}(1|s^1) = 1$, $q_{\langle S',\pi' \rangle}(0|s^{0,2}) = 1/3$ and $q_{\langle S',\pi' \rangle}(2|s^{0,2}) = 2/3$.

Let $\overline{a} \in \mathcal{A}_M$ be the task given by $\overline{a}(\theta) \coloneqq \theta$ for each $\theta \in \Theta$, and consider the monotone firm $A = \{\overline{a}\}$. This firm pays each worker their posterior expected type: $w_{\{\overline{a}\}}(s'', q, \langle S'', \pi'' \rangle) = q_{\langle S'',\pi'' \rangle}(1|s'') + 2q_{\langle S'',\pi'' \rangle}(2|s'')$ for any signal structure $\langle S'', \pi'' \rangle$ and signal $s'' \in S''$. Hence the change in average pay is

$$
\begin{aligned}
&W_{\{\overline{a}\}}(p, q, \langle S', \pi' \rangle) - W_{\{\overline{a}\}}(p, q, \langle S, \pi \rangle) \\
&= \left[ p(1) \cdot 1 + (1 - p(1)) \cdot \left( \frac{1}{3} \cdot 0 + \frac{2}{3} \cdot 2 \right) \right] - \left[ \frac{1}{4} \cdot 1 + \frac{1}{2} \cdot 2 \right] = \frac{1}{3} \left( \frac{1}{4} - p(1) \right).
\end{aligned}
$$

The instrumental component $\mathcal{I}_A(p, q, \langle S, \pi \rangle, \langle S', \pi' \rangle)$ is zero since task assignment does not change (as firm $\{\overline{a}\}$ has only one task), so the perception-correcting term is $\mathcal{C}_A(p, q, \langle S, \pi \rangle, \langle S', \pi' \rangle) = (1/4 - p(1))/3$ by Theorem 1(a).

Let $p(0) = 1/4 - 3\delta$, $p(1) = 1/4 + \delta$ and $p(2) = 1/2 + 2\delta$, where $\delta \in (-1/4, 1/12)$. If $\delta > 0$, then the population is under-perceived ($p \gtrsim_{LR} q$), but $\mathcal{C}_A(p, q, \langle S, \pi \rangle, \langle S', \pi' \rangle) < 0$. This shows that the MLR hypothesis cannot be dispensed with in Theorem 1(c). If $\delta < 0$, then the population is over-perceived ($q \gtrsim_{LR} p$), but $\mathcal{C}_A(p, q, \langle S, \pi \rangle, \langle S', \pi' \rangle) > 0$, so the MLR hypothesis cannot be dispensed with in Theorem 1(d), either.

## 4 Implications for statistical discrimination

In this section, we explore the implications of our decomposition result (Theorem 1) for pay gaps between different populations. We focus on *statistical discrimination,* meaning gaps in average pay between populations that have the same (true) skill distribution.

One message of Theorem 1 is that under some conditions, a more informative population will earn more on average. Lemma 1 says that a more favorably perceived population is paid more on average. Combining these two insights yields the following prediction about statistical discrimination:

**Corollary 2.** Consider two populations $I$ and $J$, both with skill distribution $p$ and with respective perceptions $q_I$ and $q_J$ and signal structures $\langle S_I, \pi_I \rangle$ and $\langle S_J, \pi_J \rangle$. Suppose that $\langle S_I, \pi_I \rangle$ is MLR, and that population $J$ is under-perceived ($p \gtrsim_{LR} q_J$). If population $I$ is both more favorably perceived than and more informative than population $J$ (i.e. $q_I \succsim_{LR} q_J$ and $\langle S_I, \pi_I \rangle \succsim_G \langle S_J, \pi_J \rangle$), then monotone firms pay population $I$ more on average:

$$W_A(p, q_I, \langle S_I, \pi_I \rangle) \geqslant W_A(p, q_J, \langle S_J, \pi_J \rangle) \quad \text{for any monotone firm } A \subset \mathcal{A}_M.$$

*Proof.* Fix any monotone firm $A \subset \mathcal{A}_M$. By Theorem 1(a),

$$\begin{aligned} W_A(p, q_I, \langle S_I, \pi_I \rangle) - W_A(p, q_J, \langle S_J, \pi_J \rangle) \\ &= [W_A(p, q_I, \langle S_I, \pi_I \rangle) - W_A(p, q_J, \langle S_I, \pi_I \rangle)] \\ &\quad + \mathcal{C}_A(p, q_J, \langle S_J, \pi_J \rangle, \langle S_I, \pi_I \rangle) \\ &\quad + \mathcal{I}_A(p, q_J, \langle S_J, \pi_J \rangle, \langle S_I, \pi_I \rangle). \end{aligned}$$

The bracketed term is non-negative by Lemma 1 since $q_I \succsim_{LR} q_J$. The perception-correcting "$\mathcal{C}_A$" term is non-negative by Theorem 1(c) since $p \succsim_{LR} q_J$ and since the firm $A$ is monotone and the signal structure $\langle S_I, \pi_I \rangle$ is MLR. The instrumental "$\mathcal{I}_A$" term is non-negative by Theorem 1(b). Hence $W_A(p, q_I, \langle S_I, \pi_I \rangle) - W_A(p, q_J, \langle S_J, \pi_J \rangle) \geqslant 0$. ∎

The key takeaway from Corollary 2 is that greater informativeness and more favorable perception translate straightforwardly into higher average pay *provided* the disfavored population ($J$) is under-perceived ex ante. Outside of that case, a population that is more informative and more favorably perceived may be paid *less* on average, as the following example shows.

**Example 1** (continued)**.** Consider two populations, $I$ and $J$, with equal skill type distributions $p_I = p_J = p$ and respective perceptions $q_I$ and $q_J$. Population $I$ has the fully informative signal structure $\langle S_I, \pi_I \rangle := \langle S', \pi' \rangle$, while population $J$ has the totally uninformative signal structure $\langle S_J, \pi_J \rangle := \langle S, \pi \rangle$. By Theorem 1(a), the pay gap between the two populations is

$$\begin{aligned} W_A(p, q_I, \langle S_I, \pi_I \rangle) - W_A(p, q_J, \langle S_J, \pi_J \rangle) \\ &= [W_A(p, q_I, \langle S_I, \pi_I \rangle) - W_A(p, q_J, \langle S_I, \pi_I \rangle)] \\ &\quad + \mathcal{C}_A(p, q_J, \langle S_J, \pi_J \rangle, \langle S_I, \pi_I \rangle) \\ &\quad + \mathcal{I}_A(p, q_J, \langle S_J, \pi_J \rangle, \langle S_I, \pi_I \rangle). \end{aligned}$$

By our previous calculations (see p. 11 above), the bracketed term equals $p(1) - p(1) = 0$, the perception-correcting "$\mathcal{C}_A$" term equals $p(1) - q_J(1)$, and the instrumental "$\mathcal{I}_A$" term is zero, so the total pay gap is $p(1) - q_J(1)$. Thus if population $J$ is strictly over-perceived ($q_J(1) > p(1)$), then population $I$ is paid strictly *less* than population $J$. This holds even if population $I$ is more favorably perceived ($q_I(1) \geqslant q_J(1)$).

### 4.1 The special case of accurate perceptions

Our focus in this paper is on the interplay of (mis)perceptions and information. In this section, we comment briefly on the special case of accurate perceptions, $q_I = q_J = p$, in which there is no meaningful interaction.

In particular, when perceptions are accurate, the perception-correcting effect is zero for every firm $A \subset \mathcal{A}$:

$$\mathcal{C}_A(p, p, \langle S_J, \pi_J \rangle, \langle S_I, \pi_I \rangle) \\ = \sum_{s \in S} \mu_p(s) \sum_{s' \in S'} [\mu_p(s'|s) - \mu_p(s'|s)] \sum_{\theta \in \Theta} \mu_p(\theta|s, s') \widehat{a}_s(\theta) = 0.$$

Hence by Theorem 1(a)–(b), a more informative population is always paid more when perceptions are accurate. This finding and its converse together constitute Blackwell's theorem on the value of information:[19]

**Theorem 2** (Blackwell, 1951, 1953)**.** Consider two populations $I$ and $J$, both with skill distribution $p$ and with respective perceptions $q_I$ and $q_J$ and signal structures $\langle S_I, \pi_I \rangle$ and $\langle S_J, \pi_J \rangle$. Suppose that perceptions are accurate ($q_I = q_J = p$).

(a) If $I$ is more informative than $J$ (i.e. $\langle S_I, \pi_I \rangle \succsim_G \langle S_J, \pi_J \rangle$), then every firm pays population $I$ more on average:

$$W_A(p, q_I, \langle S_I, \pi_I \rangle) \geqslant W_A(p, q_J, \langle S_J, \pi_J \rangle) \quad \text{for every firm } A \subset \mathcal{A}.$$

(b) If $I$ is not more informative than $J$ (i.e. $\langle S_I, \pi_I \rangle \not\succsim_G \langle S_J, \pi_J \rangle$), then some firm pays population $I$ strictly less on average:

$$W_A(p, q_I, \langle S_I, \pi_I \rangle) < W_A(p, q_J, \langle S_J, \pi_J \rangle) \quad \text{for some firm } A \subset \mathcal{A}.$$

Chambers and Echenique (2021) introduced the model that we employ in this paper, and used it to study statistical discrimination in the accurate-perceptions case. In particular, they sought to characterize the following strong "no-discrimination" property: "every firm $A \subset \mathcal{A}$ pays populations $I$ and $J$ the same on average, i.e. $W_A(p, p, \langle S_I, \pi_I \rangle) = W_A(p, p, \langle S_J, \pi_J \rangle)$." They obtained two geometric characterizations, proved using Choquet theory.

An alternative characterization can be obtained from Blackwell's theorem, which immediately implies that "no-discrimination" holds if and only if $\langle S_I, \pi_I \rangle \succsim_G \langle S_J, \pi_J \rangle \succsim_G \langle S_I, \pi_I \rangle$, meaning that the two populations' signal structures are informationally equivalent (i.e. they convey exactly the same information about skill).

[19] Blackwell calls skill types "states (of the world)," tasks "actions," firms "decision problems," and (expected) surplus/pay "(expected) value/payoff."

# 5 When does information narrow pay gaps?

The spirit of the perception-correcting effect is that informative signals diminish the importance for average pay of prior misperceptions. This suggests that the pay gap between two differently-perceived but equally-skilled populations narrows as the informativeness of signals improves. Bohren, Imas and Rosenberg (2019, Proposition 1) showed that this intuition is correct in their model of subjective evaluation, which can be thought of as a single-task linear–Gaussian special case of our model. (We describe their model and result in more detail in Remark 3 below.)

In this section, we explore the validity of this "information narrows pay gaps" intuition beyond the single-task linear–Gaussian special case, obtaining results that generalize and qualify the finding of Bohren, Imas and Rosenberg (2019). We focus on the case in which the two populations $I$ and $J$ have the same signal structure $\langle S, \pi \rangle$, differing only in their perceptions $q_I$ and $q_J$. Our question is under what conditions the pay gap narrows when the populations' common signal structure changes from $\langle S, \pi \rangle$ to a more informative signal structure $\langle S', \pi' \rangle$.

We identify two such conditions. The first, discussed in section 5.1, is when the change from $\langle S, \pi \rangle$ to $\langle S', \pi' \rangle$ is small. The second, discussed in section 5.2, is when $\langle S', \pi' \rangle$ is close to fully informative. Both conditions are stringent, and we show that when they are not satisfied, extra information may *widen* the pay gap. The upshot is that policy interventions or technological change which increase informativeness can either narrow or widen misperception-driven pay gaps in the labor market.

## 5.1 Slightly increased informativeness

We call a change in informativeness *slight* if it is small enough that it does not change the surplus-maximizing task assignment. The formal definition is as follows.

**Definition 2.** Fix a firm $A \subset \mathcal{A}$, a perception $q$, and signal structures $\langle S', \pi' \rangle \succsim_G \langle S, \pi \rangle$, and let $g$ be a garbling kernel from $\langle S', \pi' \rangle$ to $\langle S, \pi \rangle$. We say that $\langle S', \pi' \rangle$ is *slightly more informative than* $\langle S, \pi \rangle$ for firm $A$ at perception $q$ if

$$\left( \operatorname*{arg\,max}_{a \in A} \sum_{\theta \in \Theta} q_{\langle S, \pi \rangle}(\theta | s) a(\theta) \right) \cap \left( \operatorname*{arg\,max}_{a \in A} \sum_{\theta \in \Theta} q_{\langle S', \pi' \rangle}(\theta | s') a(\theta) \right) \neq \varnothing$$

for all signals $s \in S$ and $s' \in S'$ such that $g(s|s') > 0$.

By definition, slight increases of informativeness leave task assignment unchanged, so that the instrumental components $\mathcal{I}_A(p, q, \langle S, \pi\rangle, \langle S', \pi'\rangle)$ for $q \in \{q_I, q_J\}$ of the change in each population's average pay are zero. Note that for single-task firms $A = \{a\}$, *all* increases of informativeness are slight (whatever the perception), since such firms' task assignment is immutable.

**Remark 2.** Fix a firm $A \subset \mathcal{A}$ (a finite set, by definition) and a perception $q$. Pay as a function of the posterior belief, $r \mapsto \max_{a \in A} \sum_{\theta \in \Theta} r(\theta) a(\theta)$, is locally affine at almost every belief $r' \in \Delta(\Theta)$.[20] Hence generic signal structures $\langle S, \pi\rangle$ exclusively generate posterior beliefs at which pay is locally affine.[21] For such generic signal structures $\langle S, \pi\rangle$, any more-informative signal structure $\langle S', \pi'\rangle$ such that the posterior beliefs $q_{\langle S,\pi\rangle}(\cdot|s)$ and $q_{\langle S',\pi'\rangle}(\cdot|s')$ are always sufficiently close is slightly more informative than $\langle S, \pi\rangle$ at perception $q$.[22]

The following "positive" result identifies (admittedly stringent) conditions under which extra information does indeed narrow the pay gap.

**Proposition 1.** Fix a firm $A \subset \mathcal{A}$ and two populations $I$ and $J$, both with skill distribution $p$, signal structure $\langle S_I, \pi_I\rangle = \langle S_J, \pi_J\rangle = \langle S, \pi\rangle$, and respective perceptions $q_I$ and $q_J$. Suppose that population $I$ is more favorably perceived ($q_I \succsim_{LR} q_J$), and fix a more-informative signal structure $\langle S', \pi'\rangle$ ($\langle S', \pi'\rangle \succsim_G \langle S, \pi\rangle$). Assume that

(i) the firm $A$ is monotone ($A \subset \mathcal{A}_M$),

(ii) the signal structure $\langle S', \pi'\rangle$ is MLR,

(iii) population $I$ is over-perceived ($q_I \succsim_{LR} p$),

(iv) population $J$ is under-perceived ($p \succsim_{LR} q_J$), and

(v) $\langle S', \pi'\rangle$ is slightly more informative than $\langle S, \pi\rangle$.

Then the extra information narrows the pay gap:

$$\begin{aligned} &W_A(p, q_I, \langle S', \pi'\rangle) - W_A(p, q_J, \langle S', \pi'\rangle) \\ &\leqslant W_A(p, q_I, \langle S, \pi\rangle) - W_A(p, q_J, \langle S, \pi\rangle). \end{aligned} \tag{$\star$}$$

[20] By "locally affine at $r'$," we mean that there exists a neighborhood of $r'$ on which pay is affine. By "almost every," we mean according to the Lebesgue measure on $\Delta(\Theta)$.

[21] Formally, for any non-empty finite set $S$, it holds for (Lebesgue-)almost every $\pi : \Theta \times S \to \mathbb{R}$ such that $\langle S, \pi\rangle$ is a signal structure that for each signal $s \in S$, pay is locally affine at the posterior belief $q_{\langle S,\pi\rangle}(\cdot|s) \in \Delta(\Theta)$.

[22] Explicitly: there exists an $\varepsilon > 0$ such that if $\langle S', \pi'\rangle \succsim_G \langle S, \pi\rangle$ and $|q_{\langle S',\pi'\rangle}(\theta|s') - q_{\langle S,\pi\rangle}(\theta|s)| \leqslant \varepsilon$ for all $\theta \in \Theta$ and all pairs $(s, s') \in S \times S'$ with $\mu_q(s, s') > 0$, then $\langle S', \pi'\rangle$ is slightly more informative than $\langle S, \pi\rangle$ for firm $A$ at perception $q$.

None of the assumptions in Proposition 1 can be dispensed with: if any one of them fails, then extra information may *widen* the pay gap, as the following result shows.

**Proposition 2.** Consider tuples $(A, p, q_I, q_J, \langle S, \pi\rangle, \langle S', \pi'\rangle)$ comprising a firm $A \subset \mathcal{A}$, a skill distribution $p$, perceptions $q_I$ and $q_J$ satisfying $q_I \succsim_{LR} q_J$, and signal structures $\langle S, \pi\rangle$ and $\langle S', \pi'\rangle$ satisfying $\langle S', \pi'\rangle \succsim_G \langle S, \pi\rangle$. For each of the properties (i)–(v) in Proposition 1, there exist tuples which violate that property, satisfy all four of the other properties, and violate $(\star)$.

A key part of Proposition 2 is the assertion that slightness (property (v)) is indispensable in Proposition 1. The tuple used to prove this is as follows. Skill types are binary, $\Theta = \{0, 1\}$. The skill distribution and perceptions are $p(1) = 1/2$, $q_I(1) = 3/4$ and $q_J(1) = 1/4$, so that population $I$ is over-perceived and population $J$ is under-perceived. We consider the monotone firm $A = \{\overline{a}, \widetilde{a}\}$, where $\overline{a}(\theta) := \theta$ and $\widetilde{a}(\theta) := 4(2\theta - 1)$ for each $\theta \in \Theta$. For each $\lambda \in [1/2, 1]$, let $\langle S_\lambda, \pi_\lambda\rangle$ be the MLR signal structure given by $S_\lambda = \{s^0, s^1\}$ and $\pi_\lambda(s^0 \mid 0) = \pi_\lambda(s^1 \mid 1) = \lambda$. Note that a higher value of $\lambda$ corresponds to greater informativeness. Each population's average pay is plotted in Figure 1 as a function of $\lambda$.[23] The pay gap is not decreasing: there are $\lambda' > \lambda$ in $[1/2, 1]$ such that the pay gap is greater at $\lambda'$ than at $\lambda$.

Taken together, Propositions 1 and 2 show that whether additional information will narrow the pay gap very much depends.

**Remark 3.** In this section, where we consider only slight increases of informativeness (which leave task assignment unchanged), little insight is lost by focusing on single-task firms $A = \{a\}$. This special case is close to the model of Bohren, Imas and Rosenberg (2019), which aims to describe subjective evaluation rather than pay, and accordingly features no production choices such as task assignment. In these authors' model, it is additionally assumed that the primitives $a$, $p$, $q$ and $\langle S, \pi\rangle$ are linear–Gaussian.[24] In our language, Bohren, Imas and Rosenberg's first result implies that in the single-task linear–Gaussian case, extra information narrows the pay gap. Our Propositions 1 and 2 generalize and qualify this finding; for example, without

[23]The kinks in Figure 1 correspond to changes in task assignment. In particular, in population $I$, workers with signal $s^1$ are assigned to task $\widetilde{a}$ whatever the value of $\lambda$, while those with signal $s^0$ are assigned to $\widetilde{a}$ if $\lambda < 9/13$ and to $\overline{a}$ if $\lambda > 9/13$, and in population $J$, workers with signal $s^0$ are assigned to task $\overline{a}$ whatever the value of $\lambda$, while those with signal $s^1$ are assigned to $\overline{a}$ if $\lambda < 4/5$ and to $\widetilde{a}$ if $\lambda > 4/5$.

[24]To be precise: $\Theta = S = \mathbb{R}$, $a(\theta) = \theta$ for every $\theta \in \Theta$, and $(\theta, s) \mapsto p(\theta)\pi(s|\theta)$ and $(\theta, s) \mapsto q(\theta)\pi(s|\theta)$ are bi-variate Gaussian probability density functions.

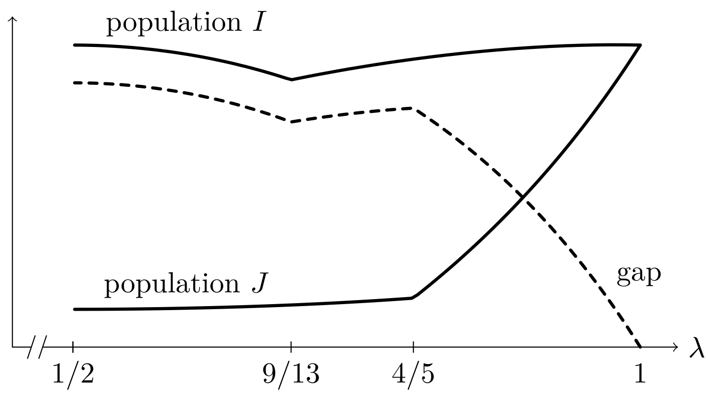


Figure 1: Average pay of populations $I$ and $J$ and the gap between them in the tuple used to prove the indispensability of property (v).

the specific linear–Gaussian structure, extra information may *widen* the pay gap unless population $I$ is over-perceived and population $J$ is under-perceived.

**Remark 1** (continued from p. 7)**.** Recall our alternative interpretation of the model, in which $q_I$ and $q_J$ are the two populations' true skill distributions, and a "like-for-like" comparison is being made between two sub-populations whose skill distributions are equal, $p_I = p_J = p$. One upshot of Propositions 1 and 2 is that extra information may narrow the pay gap in some sub-populations (those with $q_I \succsim_{LR} p \succsim_{LR} q_J$) and widen it in others (those with $p \succsim_{LR} q_I$ or $q_J \succsim_{LR} p$).

## 5.2 Near-full informativeness

Our second sufficient condition for extra information to narrow the pay gap is for the new, more informative signal structure to be "nearly fully informative," formally defined as follows.

**Definition 3.** Given $\varepsilon \geqslant 0$, we say that a signal structure $\langle S, \pi \rangle$ is *within $\varepsilon$ of full information* if for each signal $s \in S$, there exists a skill type $\theta_s \in \Theta$ such that $\pi(s|\theta) \leqslant \varepsilon \pi(s|\theta_s)$ for every other skill type $\theta \in \Theta \setminus \{\theta_s\}$.[25]

The following "positive" result shows that when informativeness is increased to near-full, the pay gap narrows.

[25]Equivalently: $\min_{\theta \in \Theta : \pi(s|\theta) > 0} [\max_{\theta' \in \Theta \setminus \{\theta\}} \pi(s|\theta') / \pi(s|\theta)] \leqslant \varepsilon$ for every $s \in S$.

**Proposition 3.** Fix a firm $A \subset \mathcal{A}$ and two populations $I$ and $J$, both with skill distribution $p$, signal structure $\langle S_I, \pi_I \rangle = \langle S_J, \pi_J \rangle = \langle S, \pi \rangle$, and respective perceptions $q_I$ and $q_J$. Suppose that population $I$ is more favorably perceived ($q_I \succsim_{LR} q_J$), and fix a more-informative signal structure $\langle S', \pi' \rangle$ ($\langle S', \pi' \rangle \succsim_G \langle S, \pi \rangle$). There exists an $\varepsilon \geqslant 0$ such that if $\langle S', \pi' \rangle$ is within $\varepsilon$ of full information, then ($\star$) on p. 17 holds. Furthermore, $\varepsilon$ may be chosen to be strictly positive, except if the right-hand side of ($\star$) is equal to zero.

Proposition 3 is a continuity result. Under a fully informative signal structure $\langle S'', \pi'' \rangle$, each signal $s'' \in S''$ perfectly reveals some skill type $\theta_{s''} \in \Theta$ (that is, $q_{\langle S'', \pi'' \rangle}(\theta_{s''} | s'') = 1$), so workers are paid

$$w_A(s'', q_I, \langle S'', \pi'' \rangle) = \max_{a \in A} a(\theta_{s''}) = w_A(s'', q_J, \langle S'', \pi'' \rangle).$$

Hence there is no pay gap: $W_A(p, q_I, \langle S'', \pi'' \rangle) - W_A(p, q_J, \langle S'', \pi'' \rangle) = 0$. Since pay as a function of the posterior belief, $r \mapsto \max_{a \in A} \sum_{\theta \in \Theta} r(\theta) a(\theta)$, is continuous, it follows that the pay gap $W_A(p, q_I, \langle S', \pi' \rangle) - W_A(p, q_J, \langle S', \pi' \rangle)$ can be made arbitrarily small (in particular, small enough that ($\star$) holds) by choosing a signal structure $\langle S', \pi' \rangle$ which produces sufficiently extreme beliefs. And the beliefs produced by signal structures $\langle S', \pi' \rangle$ that are within $\varepsilon$ of full information (in fact) become arbitrarily extreme as $\varepsilon > 0$ shrinks.

# 6 Concluding thoughts

We have studied the interplay of information and prior misperceptions in shaping average pay. Our main theoretical contribution was a decomposition of the effect of additional information into a familiar instrumental term à la Blackwell (1951, 1953) and a perception-correcting term arising from misperception. This decomposition has implications for statistical discrimination.

In order to isolate the interaction between information and misperception, we have deliberately abstracted away from other realistic frictions, such as misperceptions about signal structures (as in Bohren, Haggag, Imas & Pope, 2025), monopsony power in the labor market,[26] and agency frictions inside the firm.[27] Extending our decomposition to encompass these or other frictions is in principle straightforward: for each additional friction, there is an extra term in the decomposition capturing how new information impacts that friction, holding fixed both task assignment and the operation of the other frictions.

[26] See e.g. Boal and Ransom (1997), Manning (2003, 2021), Card (2022), Berger, Herkenhoff and Mongey (2022) and Yeh, Macaluso and Hershbein (2022).

[27] See e.g. Gibbons and Roberts (2013), Mookherjee (2006) and Holmström (2017).

# Appendix A  Proof of Lemma 1

For part (a), suppose that $q' \succsim_{LR} q$. It suffices to show that

$$w_A(s, q', \langle S, \pi\rangle) \geqslant w_A(s, q, \langle S, \pi\rangle)$$

holds for any signal structure $\langle S, \pi\rangle$, any signal $s \in S$, and any monotone firm $A \subset \mathcal{A}_M$. To that end, fix a signal structure $\langle S, \pi\rangle$ and a signal $s \in S$. Since $q' \succsim_{LR} q$, we have for any skill types $\theta'' \geqslant \theta'$ in $\Theta$ that

$$\begin{aligned} q_{\langle S,\pi\rangle}(\theta'|s) q'_{\langle S,\pi\rangle}(\theta''|s) &= \pi(s|\theta')\pi(s|\theta'')q(\theta')q'(\theta'')/k \\ &\geqslant \pi(s|\theta')\pi(s|\theta'')q(\theta'')q'(\theta')/k = q_{\langle S,\pi\rangle}(\theta''|s) q'_{\langle S,\pi\rangle}(\theta'|s) \end{aligned}$$

where $k > 0$ is a constant, which is to say that $q'_{\langle S,\pi\rangle}(\cdot|s) \succsim_{LR} q_{\langle S,\pi\rangle}(\cdot|s)$. Hence $q'_{\langle S,\pi\rangle}(\cdot|s)$ first-order stochastically dominates $q_{\langle S,\pi\rangle}(\cdot|s)$, so for any monotone firm $A \subset \mathcal{A}_M$, letting $\widehat{a}_s \in \arg\max_{a \in A} \sum_{\theta \in \Theta} q_{\langle S,\pi\rangle}(\theta|s) a(\theta)$, we have

$$\begin{aligned} w_A(s, q', \langle S, \pi\rangle) = \max_{a \in A} \sum_{\theta \in \Theta} q'_{\langle S,\pi\rangle}(\theta|s) a(\theta) &\geqslant \sum_{\theta \in \Theta} q'_{\langle S,\pi\rangle}(\theta|s) \widehat{a}_s(\theta) \\ &\geqslant \sum_{\theta \in \Theta} q_{\langle S,\pi\rangle}(\theta|s) \widehat{a}_s(\theta) = w_A(s, q, \langle S, \pi\rangle), \end{aligned}$$

where the second inequality holds since $\widehat{a}_s \in \mathcal{A}_M$ and $q'_{\langle S,\pi\rangle}(\cdot|s)$ first-order stochastically dominates $q_{\langle S,\pi\rangle}(\cdot|s)$.

For part (b), suppose that $q' \not\succsim_{LR} q$. Then there exist $\theta'' > \theta'$ in $\Theta$ such that $q'(\theta'')/q'(\theta') < q(\theta'')/q(\theta')$. Let $\langle S, \pi\rangle$ be the signal structure that sends message $s$ if skill is $\theta \in \{\theta', \theta''\}$, and otherwise perfectly reveals skill: $S = \{s\} \cup \{s_\theta\}_{\theta \in \Theta \setminus \{\theta', \theta''\}}$, $\pi(s|\theta') = \pi(s|\theta'') = 1$ and $\pi(s_\theta|\theta) = 1$ for every $\theta \in \Theta \setminus \{\theta', \theta''\}$. Fix any monotone firm $A \subset \mathcal{A}_M$. Following the imperfectly revealing signal $s \in S$, we have $q'_{\langle S,\pi\rangle}(\cdot|s) \prec_{LR} q_{\langle S,\pi\rangle}(\cdot|s)$ since

$$\frac{q'_{\langle S,\pi\rangle}(\theta''|s)}{q'_{\langle S,\pi\rangle}(\theta'|s)} = \frac{\pi(s|\theta'')}{\pi(s|\theta')} \frac{q'(\theta'')}{q'(\theta')} < \frac{\pi(s|\theta'')}{\pi(s|\theta')} \frac{q(\theta'')}{q(\theta')} = \frac{q_{\langle S,\pi\rangle}(\theta''|s)}{q_{\langle S,\pi\rangle}(\theta'|s)}.$$

Hence $q'_{\langle S,\pi\rangle}(\cdot|s)$ is strictly first-order stochastically dominated by $q_{\langle S,\pi\rangle}(\cdot|s)$, so letting $\widehat{a}'_s \in \arg\max_{a \in A} \sum_{\theta \in \Theta} q'_{\langle S,\pi\rangle}(\theta|s) a(\theta)$, we have

$$\begin{aligned} w_A(s, q', \langle S, \pi\rangle) = \sum_{\theta \in \Theta} q'_{\langle S,\pi\rangle}(\theta|s) \widehat{a}'_s(\theta) &< \sum_{\theta \in \Theta} q_{\langle S,\pi\rangle}(\theta|s) \widehat{a}'_s(\theta) \\ &\leqslant \max_{a \in A} \sum_{\theta \in \Theta} q_{\langle S,\pi\rangle}(\theta|s) a(\theta) = w_A(s, q, \langle S, \pi\rangle), \end{aligned}$$

where the strict inequality holds since $\widehat{a}'_s \in \mathcal{A}_M$ and $q'_{\langle S,\pi\rangle}(\cdot|s)$ is strictly first-order stochastically dominated by $q_{\langle S,\pi\rangle}(\cdot|s)$. When skill is fully revealed, average pay does not depend on the prior perception: for each $\theta \in \Theta \setminus \{\theta', \theta''\}$, $q'_{\langle S,\pi\rangle}(\theta|s_\theta) = q_{\langle S,\pi\rangle}(\theta|s_\theta) = 1$, so

$$w_A(s, q', \langle S, \pi\rangle) = \max_{a\in A} a(\theta) = w_A(s, q, \langle S, \pi\rangle).$$

Hence $W_A(p, q', \langle S, \pi\rangle) < W_A(p, q, \langle S, \pi\rangle)$. ■

# Appendix B Proof of Theorem 1

For part (a), define

$$K \coloneqq \sum_{s\in S} \mu_p(s) \sum_{s'\in S'} \mu_p(s'|s) \sum_{\theta\in\Theta} \mu_q(\theta|s, s')\widehat{a}_s(\theta).$$

Firstly,

$$\begin{aligned} W_A(p, q, \langle S, \pi\rangle) &= \sum_{s\in S} \mu_p(s) w_A(s, q, \langle S, \pi\rangle) \\ &= \sum_{s\in S} \mu_p(s) \sum_{\theta\in\Theta} \left[ \sum_{s'\in S'} \mu_q(s'|s)\mu_q(\theta|s, s') \right] \widehat{a}_s(\theta) \\ &= \sum_{s\in S} \mu_p(s) \sum_{s'\in S'} \mu_q(s'|s) \sum_{\theta\in\Theta} \mu_q(\theta|s, s')\widehat{a}_s(\theta) \\ &= K - \mathcal{C}_A(p, q, \langle S, \pi\rangle, \langle S', \pi'\rangle), \end{aligned}$$

where the second equality holds because $q_{\langle S,\pi\rangle}(\theta|s) = \mu_q(\theta|s)$, which equals the bracketed term by the law of total probability. Secondly,

$$\begin{aligned} W_A(p, q, \langle S', \pi'\rangle) &= \sum_{s'\in S'} \mu_p(s') w_A(s', q, \langle S', \pi'\rangle) \\ &= \sum_{s\in S} \mu_p(s) \sum_{s'\in S'} \mu_p(s'|s) \sum_{\theta\in\Theta} q_{\langle S',\pi'\rangle}(\theta|s')\widehat{a}'_{s'}(\theta) \\ &= \sum_{s\in S} \mu_p(s) \sum_{s'\in S'} \mu_p(s'|s) \sum_{\theta\in\Theta} \mu_q(\theta|s, s')\widehat{a}'_{s'}(\theta) \\ &= K + \mathcal{I}_A(p, q, \langle S, \pi\rangle, \langle S', \pi'\rangle), \end{aligned}$$

where the third equality holds since $q_{\langle S',\pi'\rangle}(\theta|s') = \mu_q(\theta|s') = \mu_q(\theta|s, s')$.

For part (b), we have for all $s \in S$ and $s' \in S'$ that

$$\sum_{\theta\in\Theta} \mu_q(\theta|s,s')\widehat{a}'_{s'}(\theta) = \sum_{\theta\in\Theta} q_{\langle S,\pi\rangle}(\theta|s')\widehat{a}'_{s'}(\theta) = \max_{a\in A} \sum_{\theta\in\Theta} q_{\langle S,\pi\rangle}(\theta|s')a(\theta)$$
$$\geqslant \sum_{\theta\in\Theta} q_{\langle S,\pi\rangle}(\theta|s')\widehat{a}_s(\theta) = \sum_{\theta\in\Theta} \mu_q(\theta|s,s')\widehat{a}_s(\theta),$$

and thus

$$\mathcal{I}_A(p,q,\langle S,\pi\rangle,\langle S',\pi'\rangle)$$
$$= \sum_{s\in S} \mu_p(s) \sum_{s'\in S'} \mu_p(s'|s) \sum_{\theta\in\Theta} \mu_q(\theta|s,s')\left[\widehat{a}'_{s'}(\theta) - \widehat{a}_s(\theta)\right] \geqslant 0.$$

For part (c), suppose in addition that $A \subset \mathcal{A}_M$, that $\langle S',\pi'\rangle$ is MLR, and that $p \gtrsim_{LR} q$. It is enough to show that

$$\sum_{s'\in S'} \left[\mu_p(s'|s) - \mu_q(s'|s)\right] \sum_{\theta\in\Theta} \mu_q(\theta|s,s')\widehat{a}_s(\theta) \geqslant 0 \quad \text{for every } s \in S.$$

To that end, we fix an arbitrary $s \in S$ and establish that

(1) $s' \mapsto \sum_{\theta\in\Theta} \mu_q(\theta|s,s')\widehat{a}_s(\theta)$ is increasing, and that

(2) $s' \mapsto \mu_p(s'|s)$ first-order stochastically dominates $s' \mapsto \mu_q(s'|s)$.

To establish claim (1), note that for any $t' > s'$ in $S'$, $\theta \mapsto \mu_q(\theta|s,t')$ is more favorable than $\theta \mapsto \mu_q(\theta|s,s')$ since for any $\theta'' > \theta'$ in $\Theta$,

$$\begin{aligned}
\mu_q(\theta'|s,s')\mu_q(\theta''|s,t') &= q_{\langle S',\pi'\rangle}(\theta'|s')q_{\langle S',\pi'\rangle}(\theta''|t') \\
&= \pi'(s'|\theta')\pi'(t'|\theta'')q(\theta')q(\theta'')/k \\
&\geqslant \pi'(t'|\theta')\pi'(s'|\theta'')q(\theta')q(\theta'')/k \\
&= q_{\langle S',\pi'\rangle}(\theta'|t')q_{\langle S',\pi'\rangle}(\theta''|s') = \mu_q(\theta'|s,t')\mu_q(\theta''|s,s')
\end{aligned}$$

for a constant $k > 0$, where the first and last equalities use the fact that $\theta$ is independent of $s$ conditional on $s'$ under $\mu_q$, and the inequality holds since $\langle S',\pi'\rangle$ is MLR. Hence $\theta \mapsto \mu_q(\theta|s,t')$ first-order stochastically dominates $\theta \mapsto \mu_q(\theta|s,s')$. Since $\widehat{a}_s \in \mathcal{A}_M$, claim (1) follows.

To establish claim (2), use bars to denote CDFs, in particular

$$\overline{\mu_r}(s'|s) \coloneqq \sum_{t'\in S' : t' \leqslant s'} \mu_r(t'|s) \quad \text{and} \quad \overline{\pi'}(s'|\theta) \coloneqq \sum_{t'\in S' : t' \leqslant s'} \pi'(t'|\theta)$$

for all $r \in \{p, q\}$, $s' \in S'$ and $\theta \in \Theta$. For each $s' \in S'$, we have

$$\mu_p(s'|s) = \sum_{\theta \in \Theta} \mu_p(\theta|s)\mu_p(s'|\theta, s)$$
$$= \sum_{\theta \in \Theta} \mu_p(\theta|s)\mu_p(s'|\theta) = \sum_{\theta \in \Theta} p_{\langle S,\pi\rangle}(\theta|s)\pi'(s'|\theta),$$

where the second equality holds since $\theta$ is independent of $s$ conditional on $s'$ under $\mu_p$. Performing the same calculation for $\mu_q$ and subtracting yields

$$\mu_p(s'|s) - \mu_q(s'|s) = \sum_{\theta \in \Theta} \Big[p_{\langle S,\pi\rangle}(\theta|s) - q_{\langle S,\pi\rangle}(\theta|s)\Big]\pi'(s'|\theta) \quad \text{for each } s' \in S',$$

and thus

$$\overline{\mu_p}(s'|s) - \overline{\mu_q}(s'|s) = \sum_{\theta \in \Theta} \Big[p_{\langle S,\pi\rangle}(\theta|s) - q_{\langle S,\pi\rangle}(\theta|s)\Big]\overline{\pi'}(s'|\theta) \quad \text{for each } s' \in S'.$$

Since $\langle S', \pi' \rangle$ is MLR, $\pi'(\cdot|\theta')$ is more favorable than $\pi'(\cdot|\theta)$ whenever $\theta' > \theta$, so $\pi'(\cdot|\theta')$ first-order stochastically dominates $\pi'(\cdot|\theta)$ whenever $\theta' > \theta$, which is to say that for each $s' \in S'$, $\theta \mapsto \overline{\pi'}(s'|\theta)$ is decreasing. As shown in the proof of Lemma 1, the fact that $p \succsim_{LR} q$ implies that $p_{\langle S,\pi\rangle}(\cdot|s)$ is more favorable than $q_{\langle S,\pi\rangle}(\cdot|s)$ for every $s \in S$, which in turn implies that for each $s \in S$, $p_{\langle S,\pi\rangle}(\cdot|s)$ first-order stochastically dominates $q_{\langle S,\pi\rangle}(\cdot|s)$. Combining these observations yields $\overline{\mu_p}(s'|s) \leqslant \overline{\mu_q}(s'|s)$ for each $s' \in S'$, which establishes claim (2).

Finally, for part (d), suppose that $q \succsim_{LR} p$. Analogously to part (c), it suffices to fix an arbitrary $s \in S$ and to establish claim (1) above and

(2′) that $s' \mapsto \mu_q(s'|s)$ first-order stochastically dominates $s' \mapsto \mu_p(s'|s)$.

Claim (1) follows from exactly the argument above, while claim (2′) follows from applying the above argument for claim (2) except with $p \succsim_{LR} q$ replaced by $q \succsim_{LR} p$. ∎

# Appendix C Proof of Proposition 1

By definition of "slightly more informative than," assumption (v) implies that each population's instrumental component is zero:

$$\mathcal{I}_A(p, q_I, \langle S, \pi\rangle, \langle S', \pi'\rangle) = \mathcal{I}_A(p, q_J, \langle S, \pi\rangle, \langle S', \pi'\rangle) = 0.$$

Hence

$$\begin{aligned}
&[W_A(p, q_I, \langle S', \pi' \rangle) - W_A(p, q_J, \langle S', \pi' \rangle)] \\
&\quad - [W_A(p, q_I, \langle S, \pi \rangle) - W_A(p, q_J, \langle S, \pi \rangle)] \\
&\qquad = \mathcal{C}_A(p, q_I, \langle S, \pi \rangle, \langle S', \pi' \rangle) - \mathcal{C}_A(p, q_J, \langle S, \pi \rangle, \langle S', \pi' \rangle) \leqslant 0,
\end{aligned}$$

where the equality holds by Theorem 1(a), and the inequality holds because the first "$\mathcal{C}_A$" term is non-positive by Theorem 1(d) (applicable by assumptions (i)–(iii)) and the second "$\mathcal{C}_A$" term is non-negative by Theorem 1(c) (applicable by assumptions (i)–(ii) and (iv)). ■

## Appendix D Proof of Proposition 2

For property (i), return to Example 2, where property (i) fails, and properties (ii) and (v) are satisfied (recall that for a single-task firm, every increase of informativeness is slight). Recall that given a skill distribution $p$ and perception $q$, the change in average pay is

$$W_{\{\underline{a}\}}(p, q, \langle S', \pi' \rangle) - W_{\{\underline{a}\}}(p, q, \langle S, \pi \rangle) = p(0) - q(0).$$

Thus if $q_I(0) < p(0) < q_J(0)$, then $(\star)$ fails, and properties (iii) and (iv) hold.

For property (ii), return to Example 3, where property (ii) fails, and properties (i) and (v) are satisfied (again recall that for a single-task firm, every increase of informativeness is slight). Given a skill distribution $p$ and perception $q$, the change in average pay is

$$\begin{aligned}
&W_{\{\overline{a}\}}(p, q, \langle S', \pi' \rangle) - W_{\{\overline{a}\}}(p, q, \langle S, \pi \rangle) \\
&\qquad = \left[p(1) \cdot 1 + (1 - p(1)) \cdot \left(\frac{q(0)}{q(0) + q(2)} \cdot 0 + \frac{q(2)}{q(0) + q(2)} \cdot 2\right)\right] \\
&\qquad\quad - \left[\frac{1}{4} \cdot 1 + \frac{1}{2} \cdot 2\right] \\
&\qquad = p(1) - \frac{5}{4} + 2(1 - p(1)) \frac{q(2)}{q(0) + q(2)},
\end{aligned}$$

so $(\star)$ fails if $p(1) < 1$ and $q_I(2)/(q_I(0) + q_I(2)) > q_J(2)/(q_J(0) + q_J(2))$. Let

$$(q_J(\theta), p(\theta), q_I(\theta)) := \begin{cases} \left(\frac{1}{4}, \frac{1}{4} - 3\delta, \frac{1}{4} - 6\delta\right) & \text{for } \theta = 0 \\ \left(\frac{1}{4}, \frac{1}{4} + \delta, \frac{1}{4} + 2\delta\right) & \text{for } \theta = 1 \\ \left(\frac{1}{2}, \frac{1}{2} + 2\delta, \frac{1}{2} + 4\delta\right) & \text{for } \theta = 2 \end{cases}$$

where $\delta = 1/25$; then ($\star$) fails, and properties (iii) and (iv) both hold.

For property (iii), suppose that skill types are binary, $\Theta = \{0, 1\}$. Let $\overline{a} \in \mathcal{A}_M$ be the task given by $\overline{a}(\theta) \coloneqq \theta$ for each $\theta \in \Theta$, and consider the firm $A = \{\overline{a}\}$; then properties (i) and (v) hold (again recall that for a single-task firm, every increase of informativeness is slight). Let $\langle S, \pi \rangle$ be an uninformative signal structure, i.e. one such that $\pi(s|\theta) = \pi(s|\theta')$ for every signal $s \in S$ and all skill types $\theta, \theta' \in \Theta$. Let $\langle S', \pi' \rangle$ be the signal structure given by $S' = \{s^0, s^1\}$ and $\pi'(s^0 \mid 0) = \pi'(s^1 \mid 1) = 3/4$, and note that property (ii) holds. Given a skill distribution $p$ and perception $q$, the change in average pay is

$$
\begin{aligned}
&W_{\{\overline{a}\}}(p, q, \langle S', \pi' \rangle) - W_{\{\overline{a}\}}(p, q, \langle S, \pi \rangle) \\
&\quad = \frac{(1 - p(1))\frac{3}{4} + p(1)\frac{1}{4}}{(1 - q(1))\frac{3}{4} + q(1)\frac{1}{4}} q(1)\frac{1}{4} + \frac{(1 - p(1))\frac{1}{4} + p(1)\frac{3}{4}}{(1 - q(1))\frac{1}{4} + q(1)\frac{3}{4}} q(1)\frac{3}{4} - q(1) \\
&\quad = \left[ \frac{3 - 2p(1)}{3 - 2q(1)} + 3\frac{1 + 2p(1)}{1 + 2q(1)} - 4 \right] \frac{q(1)}{4}.
\end{aligned}
$$

Let $p(1) = 3/4$, $q_I(1) = 1/4$ and $q_J(1) = 1/6$; then property (iii) fails, property (iv) holds, and ($\star$) fails since

$$
\begin{aligned}
&\left[ W_{\{\overline{a}\}}(p, q_I, \langle S', \pi' \rangle) - W_{\{\overline{a}\}}(p, q_I, \langle S, \pi \rangle) \right] \\
&\qquad - \left[ W_{\{\overline{a}\}}(p, q_J, \langle S', \pi' \rangle) - W_{\{\overline{a}\}}(p, q_J, \langle S, \pi \rangle) \right] = \frac{1}{10} - \frac{35}{384} > 0.
\end{aligned}
$$

For property (iv), apply the same argument as for (iii), except with $p(1) = 1/4$, $q_J(1) = 3/4$ and $q_I(1) = 5/6$.

For property (v), suppose that skill types are binary, $\Theta = \{0, 1\}$, and let the skill distribution $p$ satisfy $p(0) = p(1) = 1/2$. Define $\overline{a}, \widetilde{a} \in \mathcal{A}$ by $\overline{a}(\theta) \coloneqq \theta$ and $\widetilde{a}(\theta) \coloneqq 4(2\theta - 1)$ for each $\theta \in \Theta$, and consider the firm $A = \{\overline{a}, \widetilde{a}\}$, noting that property (i) holds. For each $\lambda \in [1/2, 1]$, let $\langle S_\lambda, \pi_\lambda \rangle$ be the MLR signal structure given by $S_\lambda = \{s^0, s^1\}$ and $\pi_\lambda(s^0 \mid 0) = \pi_\lambda(s^1 \mid 1) = \lambda$. For any perception $q$ and any $\lambda \in [1/2, 1]$, average pay at firm $A$ is

$$
\begin{aligned}
W_{\{a_1, a_2\}}(p, q, \langle S_\lambda, \pi_\lambda \rangle) &= \left[ \frac{1}{2}\lambda + \frac{1}{2}(1 - \lambda) \right] \max\left\{ r^0_{q,\lambda}, 4\left(2r^0_{q,\lambda} - 1\right) \right\} \\
&\qquad + \left[ \frac{1}{2}(1 - \lambda) + \frac{1}{2}\lambda \right] \max\left\{ r^1_{q,\lambda}, 4\left(2r^1_{q,\lambda} - 1\right) \right\} \\
&= \frac{1}{2} \max\left\{ r^0_{q,\lambda}, 4\left(2r^0_{q,\lambda} - 1\right) \right\} + \frac{1}{2} \max\left\{ r^1_{q,\lambda}, 4\left(2r^1_{q,\lambda} - 1\right) \right\}
\end{aligned}
$$

where

$$r^0_{q,\lambda} \coloneqq \frac{q(1)(1-\lambda)}{(1-q(1))\lambda + q(1)(1-\lambda)} \quad \text{and} \quad r^1_{q,\lambda} \coloneqq \frac{q(1)\lambda}{(1-q(1))(1-\lambda) + q(1)\lambda}.$$

Thus if perceptions are $q_I(1) = 3/4$ and $q_J(1) = 1/4$, and signal structures are $\langle S, \pi \rangle = \langle S_{9/13}, \pi_{9/13} \rangle$ and $\langle S', \pi' \rangle = \langle S_{4/5}, \pi_{4/5} \rangle$, then properties (ii), (iii) and (iv) are satisfied, $\langle S', \pi' \rangle$ is more informative than $\langle S, \pi \rangle$, and ($\star$) fails by direct computation (as illustrated in Figure 1). By Proposition 1, since properties (i)–(iv) hold and ($\star$) fails, it must be that property (v) fails. ■

## Appendix E Proof of Proposition 3

If a signal structure $\langle S'', \pi'' \rangle$ is fully informative, meaning that for each signal $s'' \in S''$ there is a skill type $\theta_{s''} \in \Theta$ such that $\pi''(s''|\theta_{s''}) = 1$, then

$$w_A(s'', q_I, \langle S'', \pi'' \rangle) = \max_{a \in A} a(\theta) = w_A(s'', q_J, \langle S'', \pi'' \rangle) \quad \text{for every } s'' \in S'',$$

so average pay is equal: $W_A(p, q_I, \langle S'', \pi'' \rangle) - W_A(p, q_J, \langle S'', \pi'' \rangle) = 0$.

Write $\eta \coloneqq W_A(p, q_I, \langle S, \pi \rangle) - W_A(p, q_J, \langle S, \pi \rangle)$. Since $q_I \succsim_{LR} q_J$, we have $\eta \geqslant 0$ by Lemma 1. In case $\eta = 0$, let $\varepsilon = 0$; then any signal structure $\langle S', \pi' \rangle$ that is within $\varepsilon$ of full information is itself fully informative, so $W_A(p, q_I, \langle S', \pi' \rangle) - W_A(p, q_J, \langle S', \pi' \rangle) = 0 = \eta$.

Assume for the remainder that $\eta > 0$. For any $\delta \in [0, 1]$, call a signal structure $\langle S', \pi' \rangle$ $\delta$-*extreme* if $q_{K,\langle S', \pi' \rangle}(\theta|s') \in [0, \delta] \cup [1-\delta, 1]$ for both populations $K \in \{I, J\}$, every signal $s' \in S'$, and every skill type $\theta \in \Theta$. By inspection, pay as a function of the posterior belief, $r \mapsto \max_{a \in A} \sum_{\theta \in \Theta} r(\theta) a(\theta)$, is continuous. Since any fully informative signal structure $\langle S'', \pi'' \rangle$ is 0-extreme and satisfies $W_A(p, q_I, \langle S'', \pi'' \rangle) - W_A(p, q_J, \langle S'', \pi'' \rangle) = 0$, it follows that there exists a $\delta \in (0, 1]$ such that any $\delta$-extreme signal structure $\langle S', \pi' \rangle$ satisfies $W_A(p, q_I, \langle S', \pi' \rangle) - W_A(p, q_J, \langle S', \pi' \rangle) \leqslant \eta$.

It remains only to show that for any $\delta \in (0, 1]$, there exists an $\varepsilon > 0$ such that any signal structure $\langle S', \pi' \rangle$ that is within $\varepsilon$ of full information is $\delta$-extreme. To that end, fix a $\delta \in (0, 1]$ and a $K \in \{I, J\}$, and define

$$\varepsilon \coloneqq \frac{\delta}{1-\delta} \frac{q_K(\theta_{s'})}{1 - q_K(\theta_{s'})} \frac{1}{|\Theta| - 1}.$$

$\varepsilon$ is well-defined and strictly positive since $q_K$ has full support and $|\Theta| \geqslant 2$. Fix a signal structure $\langle S', \pi' \rangle$ that is within $\varepsilon$ of full information. Then for each $s' \in S'$, there exists a skill type $\theta_{s'} \in \Theta$ such that $\pi'(s'|\theta) \leqslant \varepsilon\pi'(s'|\theta_{s'})$ for every other skill type $\theta \in \Theta \setminus \{\theta_{s'}\}$. It follows that

$$q_{K,\langle S',\pi'\rangle}(\theta_{s'}|s') = \frac{q_K(\theta_{s'})}{q_K(\theta_{s'}) + \sum_{\theta\in\Theta\setminus\{\theta_{s'}\}} \frac{\pi'(s'|\theta)}{\pi'(s'|\theta_{s'})} q_K(\theta)}$$
$$\geqslant \frac{q_K(\theta_{s'})}{q_K(\theta_{s'}) + (1-q_K(\theta_{s'}))(|\Theta|-1)\varepsilon} = 1-\delta$$

and (thus) $q_{K,\langle S',\pi'\rangle}(\theta|s') \leqslant 1 - q_{K,\langle S',\pi'\rangle}(\theta_{s'}|s') \leqslant \delta$ for every $\theta \in \Theta\setminus\{\theta_{s'}\}$. ■

# References


Agan, A., & Starr, S. (2018). Ban the box, criminal records, and racial discrimination: A field experiment. *Quarterly Journal of Economics*, *133*(1), 191–235. https://doi.org/10.1093/qje/qjx028

Aigner, D. J., & Cain, G. G. (1977). Statistical theories of discrimination in labor markets. *Industrial and Labor Relations Review*, *30*(2), 175–187. https://doi.org/10.1177/001979397703000204

Alonso, R., & Câmara, O. (2016). Bayesian persuasion with heterogeneous priors. *Journal of Economic Theory*, *165*, 672–706. https://doi.org/10.1016/j.jet.2016.07.006

Arnold, D., Dobbie, W., & Yang, C. S. (2018). Racial bias in bail decisions. *Quarterly Journal of Economics*, *133*(4), 1885–1932. https://doi.org/10.1093/qje/qjy012

Arrow, K. J. (1973). The theory of discrimination. In O. Ashenfelter & A. Rees (Eds.), *Discrimination in labor markets* (pp. 3–33). Princeton University Press. https://doi.org/10.1515/9781400867066-003

Azmat, G., & Petrongolo, B. (2016). Gender and the labor market: What have we learned from field and lab experiments? *Labour Economics*, *30*, 32–40. https://doi.org/10.1016/j.labeco.2014.06.005

Becker, G. S. (1955). *The economics of racial discrimination* [Doctoral dissertation, University of Chicago].

Becker, G. S. (1957). *The economics of discrimination*. University of Chicago Press.

Berger, D., Herkenhoff, K., & Mongey, S. (2022). Labor market power. *American Economic Review*, *112*(4), 1147–93. https://doi.org/10.1257/aer.20191521

Bergson, A. (1938). A reformulation of certain aspects of welfare economics. *Quarterly Journal of Economics*, *52*(2), 310–334. https://doi.org/10.2307/1881737

Bertrand, M., & Duflo, E. (2017). Field experiments on discrimination. In A. V. Banerjee & E. Duflo (Eds.), *Handbook of economic field experiments* (pp. 309–393, Vol. 1). North-Holland. https://doi.org/10.1016/bs.hefe.2016.08.004

Blackwell, D. (1951). Comparison of experiments. In J. Neyman (Ed.), *Berkeley symposium on mathematical statistics and probability* (pp. 93–102, Vol. 2). University of California Press.

Blackwell, D. (1953). Equivalent comparisons of experiments. *Annals of Mathematics and Statistics*, *24*(2), 265–272. https://doi.org/10.1214/aoms/1177729032

Blackwell, D., & Dubins, L. (1962). Merging of opinions with increasing information. *Annals of Mathematical Statistics*, *33*(3), 882–886. https://doi.org/10.1214/aoms/1177704456

Blau, F. D., & Kahn, L. M. (2017). The gender wage gap: Extent, trends, and explanations. *Journal of Economic Literature*, *55*(3), 789–865. https://doi.org/10.1257/jel.20160995

Boal, W. M., & Ransom, M. R. (1997). Monopsony in the labor market. *Journal of Economic Literature*, *35*(1), 86–112.

Bohren, J. A., Haggag, K., Imas, A., & Pope, D. G. (2025). Inaccurate statistical discrimination: An identification problem. *Review of Economics and Statistics*, *107*(3), 605–620. https://doi.org/10.1162/rest_a_01367

Bohren, J. A., Imas, A., & Rosenberg, M. (2019). The dynamics of discrimination: Theory and evidence. *American Economic Review*, *109*(10), 3395–3436. https://doi.org/10.1257/aer.20171829

Bordoli, D. (2025). *Non-Bayesian updating and value of information* [working paper, 1 Jun 2025]. https://doi.org/10.2139/ssrn.5280727

Card, D. (2022). Who set *your* wage? *American Economic Review*, *112*(4), 1075–1090. https://doi.org/10.1257/aer.112.4.1075

Chambers, C. P., & Echenique, F. (2021). A characterisation of 'Phelpsian' statistical discrimination. *Economic Journal*, *131*(637), 2018–2032. https://doi.org/10.1093/ej/ueaa107

Dworczak, P., Kominers, S. D., & Akbarpour, M. (2021). Redistribution through markets. *Econometrica*, *89*(4), 1665–1698. https://doi.org/10.3982/ECTA16671

Fang, H., & Moro, A. (2011). Theories of statistical discrimination and affirmative action: A survey. In J. Benhabib, M. O. Jackson, & A. Bisin (Eds.), *Handbook of social economics* (pp. 133–200, Vol. 1A). North-Holland. https://doi.org/10.1016/B978-0-444-53187-2.00005-X

Gibbons, R., & Roberts, J. (Eds.). (2013). *The handbook of organizational economics*. Princeton University Press. https://doi.org/10.1515/9781400845354

Girsanov, I. V. (1960). On transforming a certain class of stochastic processes by absolutely continuous substitution of measures (E. Nishiura, Trans.). *Theory of Probability and its Applications*, *5*(3), 285–301. https://doi.org/10.1137/1105027

Holmström, B. (2017). Pay for performance and beyond. *American Economic Review*, *107*(7), 1753–77. https://doi.org/10.1257/aer.107.7.1753

Kartik, N., Lee, F. X., & Suen, W. (2021). Information validates the prior: A theorem on Bayesian updating and applications. *American Economic Review: Insights*, *3*(2), 165–822. https://doi.org/110.1257/aeri.20200284

Lang, K., & Lehmann, J.-Y. K. (2012). Racial discrimination in the labor market: Theory and empirics. *Journal of Economic Literature*, *50*(4), 959–1006. https://doi.org/10.1257/jel.50.4.959

Lang, K., & Spitzer, A. K.-L. (2020). Race discrimination: An economic perspective. *Journal of Economic Perspectives*, *34*(2), 68–89. https://doi.org/10.1257/jep.34.2.68

Laplace, P.-S. (1774). Mémoire sur la probabilité des causes par les évènements. *Mémoires de Mathématique et de Physique, Présentés à l'Académie Royale des Sciences*, *6*, 621–656.

Manning, A. (2003). *Monopsony in motion: Imperfect competition in labor markets*. Princeton University Press. https://doi.org/10.2307/j.ctt5hhpvk

Manning, A. (2021). Monopsony in labor markets: A review. *ILR Review*, *74*(1), 3–26. https://doi.org/10.1177/0019793920922499

Mookherjee, D. (2006). Decentralization, hierarchies, and incentives: A mechanism design perspective. *Journal of Economic Literature*, *44*(2), 367–390. https://doi.org/10.1257/jel.44.2.367

Neumark, D. (2018). Experimental research on labor market discrimination. *Journal of Economic Literature*, *56*(3), 799–866. https://doi.org/10.1257/jel.20161309

Office for National Statistics. (2023). *Ethnicity pay gaps, UK: 2012 to 2022*. Retrieved August 18, 2025, from https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/earningsandworkinghours/articles/ethnicitypaygapsingreatbritain/2012to2022

Office for National Statistics. (2024a). *Disability pay gaps, UK: 2014 to 2023*. Retrieved August 18, 2025, from https://www.ons.gov.uk/peoplepop ulationandcommunity/healthandsocialcare/disability/articles/disabilitypaygapsintheuk/2014to2023

Office for National Statistics. (2024b). *Gender pay gap in the UK: 2024*. Retrieved August 18, 2025, from https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/earningsandworkinghours/bulletins/genderpaygapintheuk/2024

Onuchic, P. (2023). *Recent contributions to theories of discrimination* [working paper, 1 Jun 2023]. https://doi.org/10.48550/arXiv.2205.05994

Onuchic, P., & Ray, D. (2023). Conveying value via categories. *Theoretical Economics*, *18*(4), 1407–1439. https://doi.org/10.3982/TE5026

Phelps, E. S. (1972a). *Inflation policy and unemployment theory: The cost–benefit approach to monetary planning*. Macmillan.

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

Rothstein, J. M. (2004). College performance predictions and the SAT. *Journal of Econometrics*, *121*(1–2), 297–317. https://doi.org/10.1016/j.jeconom.2003.10.003

Samuelson, P. A. (1947). *Foundations of economic analysis*. Harvard University Press.

U.S. Department of Labor. (2025). *Earnings and earnings ratios by sex, race, and occupation group*. Retrieved August 18, 2025, from https://www.dol.gov/agencies/wb/data/earnings/wage-gap-race-occupation

Wald, A. (1947). *Sequential analysis*. Wiley.

Yeh, C., Macaluso, C., & Hershbein, B. (2022). Monopsony in the US labor market. *American Economic Review*, *112*(7), 2099–2138. https://doi.org/10.1257/aer.20200025