An information theory of efficient differential treatment

Emil Temnyalov*

November 1, 2021.

Abstract

When are differential treatment policies—such as preferential treatment, affirmative action, and gender equity policies—justified by efficiency concerns? I propose a non-parametric assignment model where a policymaker assigns agents to different treatments or positions to maximize total surplus, based on the agents' characteristics and on noisy information about their types. I provide necessary and sufficient conditions on the agents' signal structures which characterize whether surplus maximization requires differential treatment or not, and study how the bias and informativeness of signal structures determine the efficiency implications of differential treatment. I examine implications of this model for inequality, decentralization and empirical work.

Keywords: optimal assignment, signaling, efficiency, differential treatment, affirmative action, gender equity, education, labor markets. **JEL codes:** D47, D61, D8, J7, I24.

*University of Technology Sydney, Department of Economics. emil.temnyalov@uts.edu.au.

This paper was previously titled "An economic theory of differential treatment." This research was funded partially by the Australian Government through the Australian Research Council's Discovery Early Career Researcher Award: DE200100590. I thank the Coeditor, Leslie Marx, and 3 anonymous referees for their detailed and constructive comments. For helpful comments I also thank Dan Bernhard, Juan Carlos Carbajal, Vessela Daskalova, Federico Echenique, Jeffrey Ely, Isa Hafalir, Young-Chul Kim, Bettina Klaus, Hao Li, George Mailath, Thomas Mariotti, Shiko Maruyama, Alessandro Pavan, Bruno Strulovici, Kentaro Tomoeda, Steven Williams, and seminar participants at Caltech, Collegio Carlo Alberto, Simon Fraser, Toulouse School of Economics, UBC, UC Dublin, UIUC, UNSW, as well as conference participants at AETW 2019, ES NASM 2019, PET 2019, OEW 2019, APIOC 2019, AEA 2020, EEA 2020, CMID 2020, MD4SG 2020, ES World Congress 2020, EAAMO 2021, and the GTS World Congress 2020/21.

1 Introduction

Discriminatory policies are often controversial in debates about public policy, education, hiring and promotion within organizations, and in other contexts with differential treatment, whereby economic agents are treated differently depending on some observable characteristics. In an education context, for example, whether a student is admitted to a university or not may depend on their socio-economic status, race, or some other characteristic of them or their environment.[1] Similarly, in a labor market context, an organization's promotion or hiring decisions may depend not just on the performance of a candidate, but also on the opportunities that the person has had, their gender, or some other characteristic of them or their environment.[2]

Debates about differential treatment policies are often based on concerns about fairness, and one can make arguments both in favor of and against differential treatment. While recognizing that fairness is an important ethical criterion for evaluating such policies, this paper takes a different but complementary approach, and considers differential treatment from the perspective of efficiency rather than fairness. Recent work has shown that diversity may improve efficiency in many settings, including in organizations (Herring, 2009), in marketplaces (Levine et al., 2014), in innovation and entrepreneurship (Nathan and Lee, 2013), in teamwork (Hoogendoorn et al., 2013), in education (Hansen et al., 2015), and at the country level (Alesina et al., 2016). I propose that efficiency may inherently require differential treatment because of information frictions. I show this in a general and parsimonious model with heterogeneous signaling technologies, which does not feature any intrinsic benefits from differential treatment, such as benefits from diversity or from targeting agents with special characteristics.

I study an assignment problem where a policy-maker must assign agents to different treatments or positions, based on some observable signals and characteristics of the agents. The policy-maker's objective is to design an assignment policy which maximizes total expected surplus, subject to an ex post feasibility constraint—i.e. the number of agents assigned to each position cannot exceed the capacity of that position. I highlight two leading exam-

[1] For instance, the recent "Students for Fair Admissions v. Harvard" lawsuit alleged that Harvard "systematically and unconstitutionally discriminates against Asian-American applicants by penalizing their high achievements as a group, while giving preferences to other racial and ethnic minorities" (Hartocollis, 2018).

[2] Blau and Kahn (2017); Matysiak and Cukrowska-Torzewska (2021) review recent trends and explanations of the gender gap in employment and policies in this context.

ples from education and labor contexts: a university's admissions office must decide which students to admit or not; and an organization must decide which job candidates to hire, or how to promote employees to different ranks within the organizational hierarchy. Each agent generates some surplus or value (e.g. student achievement or employee output), which depends on their unobservable type (e.g. a student's ability or an employee's productivity) and on the treatment or position to which they are assigned. The policy-maker observes a noisy signal of the agent's type (e.g. a student's standardized test score or an employee's past performance), and also some characteristics or categories of the agent (e.g. a student's socio-economic status, race, gender, or an employee's demographics, gender, etc.)

This paper characterizes the surplus-maximizing policy in a general setting, under mild non-parametric assumptions on values and signal distributions. First, I show that the optimal assignment policy will generically feature differential treatment across the observable characteristics of the agents, even if these characteristics are not directly payoff relevant, as long as the distributions of signals vary across these observable characteristics. More precisely, the optimal assignment policy is assortative with respect to a *benefit index*, which measures the expected incremental benefit from being assigned to a higher position or treatment. But the optimal policy need not be monotonic with respect to signals, because signals do not necessarily translate monotonically into expected incremental benefits, if the conditional distributions of types, given signals, differ across characteristics. The optimal policy is deterministic (except for tie-breaking among payoff-equivalent agents) and can be implemented with category-specific thresholds, i.e. signal cutoffs which vary across agents' characteristics. In fact, I show that the optimal policy is only assortative with respect to signals, for any capacities, if and only if the conditional signal distributions are identical across observable characteristics in a particular sense. This is a very strong condition, and it suggests that in many realistic settings the surplus-maximizing policy will feature some differential treatment which favors agents whose signal distributions are "worse"—i.e. a form of positive discrimination.

Second, I show that the optimal policy may feature differential treatment even in cases where signals are unbiased predictors of agents' types, i.e. an agent's expected type, conditional on a signal, does not vary across other observable characteristics. This is because the policy-maker's objective function may induce a preference for or against dispersion in types, and the conditional type distributions can be more noisy or less noisy across different characteristics, even if expected types are equal across characteristics. That is, some categories of agents may have more or less informative signals, defined in terms of mean-preserving spreads in conditional type distributions, and as a result may optimally be treated differently, depending on the shape of the surplus function. Hence the optimal policy is monotonic with respect to the benefit index, but need not be monotonic with respect to the agents' expected types, and may favor agents whose type distributions are either more dispersed or less dispersed.

To derive the main results, outlined above, I provide a definition of *comparability* across categories, which requires that the expected incremental benefits from any two positions, as a function of an agent's signal, can be translated across characteristics. This notion allows signals to be compared across different characteristics, in an environment with any signal distributions. Comparability implies an ordering over type distributions, which is more general than the usual stochastic orderings found in the literature, since it imposes restrictions only on the expected values of the conditional type distributions, rather than on the shape of the distributions. Moreover, comparability is an easy property to verify in many applications, and is satisfied in virtually every parametric environment previously studied in the literature. In particular, I provide 3 different sufficient conditions for the notion of comparability to hold: either (i) treatments are binary; or (ii) the surplus is multiplicatively separable in type and position; or (iii) the signal distributions are similar in a particular sense.

My main results can be interpreted, for example, in the context of university admissions policies. One might think that the distribution of students' university exam scores, conditional on their true abilities, varies across socio-economic status or some other observable characteristic (e.g. because parents invest different amounts of time and resources in their children's education, depending on their status). A university admissions office, seeking to maximize surplus, would therefore use socio-economic status information in forming beliefs about students' true abilities. My first result shows that the optimal admissions policy may admit students non-monotonically with respect to their exam scores. A lower-score student from a socio-economic group whose score distributions are worse may be preferred over a higher-score student from a socio-economic group with higher score distributions. This highlights a *bias* motive for differential treatment. Surplus is a function of agents' types, rather than signals, so the optimal policy must interpret signals differently if their distributions are biased across categories. Furthermore, the optimal policy may favor some group even if signals are not biased (i.e. expected abilities conditional on scores are the same across groups), for example if the variance of abilities is different across characteristics. Whether more or less noisy signals result in more or less favorable treatments across groups depends on whether there are increasing or decreasing marginal values from higher types (e.g. if the value of "superstar" students is disproportionally high). This highlights an *informativeness* motive for differential treatment: the informational content of signal distributions can vary across characteristics and thus affect the dispersion of types across characteristics. It is important to note that these results apply not only to affirmative action in university admissions, but also to differential treatment in school choice, such as in the current debate about how to improve access for lower socio-economic status (SES) students to elite selective public schools in New York's and Chicago's school districts. These bias and informativeness motives can also arise in hiring and promotion decisions within organizations, for example in differential treatment based on gender.

In addition to my main results I also discuss several more specific applications of the model. First, I consider an environment with biased signals, where the distributions of types conditional on any signal can be ordered across characteristics in terms of first-order stochastic dominance. In this application signals over-estimate types for higher categories and underestimate them for lower ones. Such a model may accurately describe the relationship between student abilities, standardized test scores, and socio-economic status. The optimal policy here favors agents from categories with worse signals, and the optimal cutoffs for each position increase across categories; i.e. lower-SES students have lower signal cutoffs.

Second, I consider an environment with unbiased but noisier signals, where type distributions conditional on signals are mean-preserving spreads across categories, and the value function features a preference either for or against dispersion. The optimal policy implies that the aggregate treatments vary across categories, although the unconditional type distributions are the same a priori. Hence this environment creates endogenous aggregate inequality. For example, if the policy-maker has a preference for dispersion, then agents from categories with noisier conditional type distributions will be favored for each position, creating aggregate inequality.

Third, I compare centralized and decentralized assignment policies and study the effect of decentralization on efficiency. This analysis is relevant to a range of settings where decision-makers' objectives may be to maximize the surplus they generate within their unit, rather than the overall value of the assignment. I find that the notion of comparability plays a crucial role in this analysis: decentralized assignment is efficient if and only if categories satisfy a strong version of comparability.

2 Related literature

To my knowledge this is the first paper to study under what conditions efficient assignments require differential treatment. At a fundamental level, the idea that an optimal policy may condition not just on what is directly surplus-relevant, but also on non-surplus-relevant characteristics, is not new to economics.[3] The first contribution of this paper is to formalize this idea in a general and tractable non-parametric model of optimal assignment, which has the scope to be extended to many new differential treatment environments, including mechanism design, organizational design, and information design. The paper also identifies a particular concept of comparability as a useful property for tractability.

This paper relates to a growing literature on the effect and implementation of diversity and distributional objectives in matching and assignment models. Several papers study controlled school choice, in the form of quotas or reserves, and the effects of affirmative action, a form of differential treatment. This literature includes Abdulkadiroğlu and Sŏnmez (2003); Abdulkadiroğlu (2005); Kojima (2012); Hafalir et al. (2013); Ehlers et al. (2014). Hafalir et al. (2019) study inter-district school choice designed to balance diversity. Fu (2007) models college admissions as an asymmetric all-pay auction with complete information, where maximizing student efforts requires a handicap that resembles affirmative action. Olszewski and Siegel (2019) study college admissions with centralized tests and costly effort, and discuss how coarse grading policies can lead to Pareto improvements. Bodoh-Creed and Hickman (2018) study college admissions with vertically differentiated colleges and endogenous human capital, and establish the equivalence of quota and admission preference systems. Che and Koh (2016) model decentralized college admissions, characterize equilibria and compare them to centralized admissions. Unlike the latter papers, I develop a model with a finite number of agents, which can be applied to settings that don't suit the large-market framework. Moreover, this model explicitly distinguishes between agents' payoff-relevant types and their signals, whereas the matching literature typically interprets signals as types. The latter distinction is critical for studying the efficiency implications of differential treatment.

[3] For example, a large public finance literature on "tagging," beginning with Akerlof (1978), studies whether conditioning payments on agents' characteristics can improve the effectiveness of public finance programs. This idea has also been applied in development contexts, e.g. in the targeting of anti-poverty measures (Besley and Kanbur, 1988; Coady et al., 2004). I consider similar differential treatment policies in an assignment model that can be applied to school choice, university admissions, or hiring and promotion, where assignments are a function of signals whose informational content can vary across characteristics.

More broadly, a large economics literature studies statistical discrimination, beginning with papers by Phelps (1972) and Arrow (1973), which offered an alternative to the model of taste-based discrimination by Becker (1957), summarized in a comprehensive survey by Fang and Moro (2011). These seminal papers have created two strands of literature: one where there are exogenous differences between different groups of agents, giving rise to discrimination, and another where differences between groups of agents arise endogenously in equilibrium, even though groups are ex ante identical. In the first strand, Chambers and Echenique (2021) study the connection between statistical discrimination and the statistical identification of workers' signals from skills. Aigner and Cain (1977) and Lundberg and Startz (1983) study human capital investment in a model where different groups have more or less noisy signals of productivity. Moro and Norman (2003) study human capital in a model with heterogeneous investment costs across groups. Cornell and Welch (1996) study labor tournaments where groups have access to more or fewer signals of their productivity, and the scarcity of jobs leads to differential treatment across groups. In the second strand, Coate and Loury (1993) study noisily observed skill investments by workers, whereby some groups may endogenously be favored by employers. Moro and Norman (2004) study discrimination which arises with interactions between groups. Mailath et al. (2000) study discrimination that arises due to search frictions, even with perfect information about agents' types. In contrast, this paper focuses on designing assignment policies which maximize efficiency, rather than on explaining discrimination per se. The paper also contributes to the statistical discrimination literature by developing a more general and completely non-parametric framework. Importantly, I show that some of the findings of this literature on the effect of differential noisiness of signals can be reversed. While the literature generally finds that groups with *less noisy* signals are treated favorably, I show that the opposite can also happen. I find precise conditions on the primitive surplus function which determine whether higher signal noise leads to more or less favorable treatment. Finally, I show that aggregate inequality across groups can arise even in the absence of human capital investment, search frictions, coordination failure, or interactions across groups.

A subset of the discrimination literature considers affirmative action as a possible policy remedy to discrimination. This literature generally finds that affirmative action may exacerbate differences across groups, e.g. in terms of human capital, and may lead to lower welfare (Coate and Loury, 1993; Moro and Norman, 2003, 2004; Norman, 2003; Fang and Norman, 2006). It is important to note, however, that affirmative action is distinct from merely banning statistical discrimination, on which much of the literature focuses. I consider a richer model of differential treatment which applies to more general policies beyond affirmative action or banning or allowing statistical discrimination.

Other models of the effect of affirmative action include Chung (2000), who provides a theory of role models. Chan and Eyster (2003) study color-blind affirmative action in a university admissions model, where signals are directly payoff-relevant and the university has an explicit preference for diversity. Epple et al. (2008) study a related model of competing colleges with a preference for diversity. Fryer and Loury (2007) consider a similar model with an added ex ante investment stage, where groups have exogenous differences in skill investment costs, and analyze the optimal timing of interventions. Garg et al. (2021) study a Gaussian model of admissions where the policy-maker explicitly values diversity and consider the effect of dropping standardized tests. In contrast to this literature, the policy-maker in my model has no preference for diversity and only seeks to maximize expected total surplus. Moreover, signals are distinct from types and only the latter are directly payoff-relevant. Finally, I do not assume any a priori differences in the distributions of types across groups, and differences arise only due to signaling.

3 Model

A policy-maker wants to assign agents to treatments or positions. There is a set of agents N with cardinality n, and an ordered set of vertically differentiated positions $P = \{p_1, ..., p_m\}$ with corresponding capacities $\{k_1, ..., k_m\} > 0$, where $\sum_j k_j \geq n$. The latter is without loss of generality, as p_1 can be a "null" or "unassigned" position with unlimited capacity.

Each agent has an unknown type $t_i \in T$, drawn from a distribution F, and a publicly observed category or characteristic $x_i \in X$. The policy-maker observes a noisy signal of each agent's type, $s_i \in S$, drawn from a distribution F_{t_i,x_i}. I assume S is an ordered convex set, T is an ordered set with discrete or continuous values, and X is any set of (possibly vector-valued) characteristics with discrete or continuous values. The set of characteristics need not be ordered, e.g. when considering gender or ethnicity. The orders imposed on types and signals are natural in many applications: higher or lower types represent ability or productivity, while higher or lower signals represent test scores of students or past productivity of employees.

A policy is a stochastic assignment of agents mapping observables to positions, denoted by a function $\mathscr{P} : (X \times S)^n \to \Delta(P^n)$. I denote by $\mathbb{P}(p|x_i, s_i)$ the distribution of agent i's position, and by $p(x_i, s_i)$ the realized position. Every possible assignment generates some individual ex post surplus, given by $v : T \times P \to \mathbb{R}$, where v is increasing in both arguments.[4]

The policy-maker's objective is to design a policy that maximizes the expected total surplus from the assignment, subject to ex post feasibility:

$$\max_{\mathscr{P}} \quad \sum_i \mathbb{E}[v(t_i, p(x_i, s_i))] \quad = \quad \sum_i \sum_j \int_T v(t_i, p_j) dF(t_i|x_i, s_i) \cdot \mathbb{P}(p_j|x_i, s_i) \qquad (1)$$
$$\text{s.t.} \quad |\{i : p(x_i, s_i) = p_j\}| \leq k_j$$

I make three standard assumptions to analyze this problem.

Assumption 1 (Supermodularity). *The surplus v is supermodular (equivalently, has increasing differences) in (t, p): $v(t'', p'') + v(t', p') > v(t'', p') + v(t', p'')$ for all $t'' > t'$ and $p'' > p'$.*

Assumption 1 represents complementarity between the agent's type and their position. Supermodularity of the surplus function is sufficient to guarantee that the first-best assignment features positive assortative matching; i.e. it is optimal to assign higher positions to higher types. Such assortative matching is common in all the motivating examples for this paper, and supermodularity is commonly used to motivate this observation in many applications.

Assumption 2 (Continuity). *The expected surplus $\mathbb{E}[v(t, p)|x, s]$ is continuous in s.*

Assumption 2 is a mild technical assumption for practical applications, as it requires only that the expected surplus of an agent varies continuously in s, holding the position p and the category x constant. This assumption is satisfied whenever the density $f(t|s)$ is uniformly continuous is s, which by the monotone convergence theorem implies that $\mathbb{E}[v(t, p)|x, s]$ is continuous in s. Moreover, the main intuitions of this paper do not crucially rely on continuity, and one can state analogous results in the absence of this assumption.

[4] Note that this surplus depends on types, whereas signals are not directly payoff-relevant. Alternatively, one can incorporate endogenous human capital investment in an extension of the model where $V : S \times T \times P \to \mathbb{R}$. In such a setting signals are themselves productive and the surplus is a function not only of the agent's type, but also of an investment choice that generates the signal S. As will be clear from the subsequent analysis, my main results on differential treatment would continue to hold qualitatively in such an extension, but the degree of differential treatment will be moderated to some extent. The optimal assignment will generally feature a smaller differential compared to the case where signals are not productive. This setting offers a fruitful avenue for future research.

Assumption 3 (MLRP). *The family of signal distributions $\{F_{t,x}\}$ satisfies the strict monotone likelihood ratio property in t: for $s'' > s'$, the ratio of the densities $\frac{f_{t,x}(s'')}{f_{t,x}(s')}$ is increasing in t, conditional on x, whenever it is well-defined.*

Assumption 3 is a standard assumption in the information economics literature. It ensures that there is an intuitive way to interpret signals and update beliefs regarding an agent's type, conditional on observing some signal, as in the following proposition.

Proposition 1 (Milgrom 1981). *Under Assumption 3, $F(\cdot|x_i, s_i = s'') \succ_{FOSD} F(\cdot|x_i, s_i = s')$ for any $s'' > s'$.*

Proposition 1, due to Milgrom (1981), shows that observing a higher signal s is good news regarding the agent's type (conditional on any category x); i.e. a higher signal improves the distribution of t in the sense of first-order stochastic dominance. This proposition implies that $\mathbb{E}[v(t,p)|x,s] = \int_T v(t,p)dF(t|x,s)$ is increasing in s, for any x, p, since $v(t,p)$ is increasing in t and $F(t|x,s)$ is increasing in s in the first order. Furthermore, the proposition also implies that $\mathbb{E}[v(t,p'')|x,s] - \mathbb{E}[v(t,p')|x,s] = \int_T v(t,p'') - v(t,p')dF(t|x,s)$ is increasing in s, for any x and $p'' > p'$, because $F(t|x,s)$ is increasing in s in the first order and $v(t,p)$ is supermodular, hence $v(t,p'') - v(t,p')$ is an increasing function of t.

Comparability

The assumptions above hold within categories, so they do not impose any restrictions on how agents from different categories compare to each other. In principle, the distributions of types and signals may differ in arbitrary ways across categories. For example, as functions of x, these distributions may be shifting or spreading, or may not be comparable using any of the standard stochastic orders at all. I define two notions of comparability across different categories, which impose conditions only on the expected values of the distributions, rather than on the whole distribution functions. They allow for comparisons of distributions that need not be comparable according to FOSD, SOSD, etc., and they may also be easier to test empirically, as they only require comparing conditional expectations, rather than more complex distributional properties. I first define *strong comparability*, which requires that agents with different signals can be compared across categories in terms of their absolute levels of expected surplus. Then I define *comparability*, a weaker notion which requires that agents with different signals can be compared across categories in terms of their relative gains in expected surplus. The former notion is not necessary for my main results, but it relates to several of the propositions in section 3.3 and is easier to state.

Definition 1. *Categories x, x' are **strongly comparable** if $\forall s \in S$ one of the following holds:*

(i) $\exists s' \in S$ *s.t.* $\mathbb{E}[v(t,p)|x,s] = \mathbb{E}[v(t,p)|x',s'] \, \forall p$;
(ii) $\forall s' \in S$, $\mathbb{E}[v(t,p)|x,s] \geq \mathbb{E}[v(t,p)|x',s'] \, \forall p$;
(iii) $\forall s' \in S$, $\mathbb{E}[v(t,p)|x,s] \leq \mathbb{E}[v(t,p)|x',s'] \, \forall p$.

This definition establishes an equivalence across categories, defined in terms of expected surplus as a function of signals. Strong comparability requires that an agent with characteristics x and a signal s generates the same expected surplus in any position p as an agent with characteristics x' and *some* (possibly different) signal s'. Hence agents with (x, s) and (x', s') are strongly comparable because they have the same expected surplus across all positions. In this sense one can "translate" or map signals s and s' across the two categories x and x', and this translation is invariant across positions. The second and third cases of the definition also allow for the possibility that the (x, s) agent generates an expected surplus that is either uniformly higher or uniformly lower across all positions than that of agents of category x' with *any* signal. Therefore (x, s) can either be mapped to a particular (x', s'), or has higher expected surplus than all (x', s') or lower surplus than all (x', s').

The next definition provides a less demanding notion of comparability, which is implied by strong comparability and which makes the main results of the paper applicable to a wider set of environments.

Definition 2. *Categories x, x' are **comparable** if $\forall s \in S$ one of the following holds:*

(i) $\exists s' \in S$ *s.t.* $\mathbb{E}[v(t,p)|x,s] - \mathbb{E}[v(t,p')|x,s] = \mathbb{E}[v(t,p)|x',s'] - \mathbb{E}[v(t,p')|x',s'] \, \forall p, p'$;
(ii) $\forall s' \in S$, $\forall p$ *and* $p' < p$, $\mathbb{E}[v(t,p)|x,s] - \mathbb{E}[v(t,p')|x,s] \geq \mathbb{E}[v(t,p)|x',s'] - \mathbb{E}[v(t,p')|x',s']$;
(iii) $\forall s' \in S$, $\forall p$ *and* $p' < p$, $\mathbb{E}[v(t,p)|x,s] - \mathbb{E}[v(t,p')|x,s] \leq \mathbb{E}[v(t,p)|x',s'] - \mathbb{E}[v(t,p')|x',s']$.

This weaker notion requires that only the *relative gains* in expected surplus across different positions be comparable, rather than the *absolute levels* of expected surplus. That is, for two categories to be comparable, agents need to have the same expected gain from being assigned to a position p instead of p', for any $p > p'$. Analogously to strong comparability, the second and third cases of the definition allow for the possibility that the (x, s) agent has relative gains that are either uniformly higher or uniformly lower across all positions than those of agents of category x' with *any* signal. Under comparability (x, s) can either be mapped to a particular (x', s'), or has larger relative gains in expected surplus than all (x', s') or smaller gains than all (x', s').

To build some intuition, one can interpret these definitions in the context of school choice. Consider a district with a selective and a non-selective school. Strong comparability means that a student with characteristics x and signal s has the same expected returns to education from both the selective and the non-selective school as a student with characteristics x' and some (possibly different) signal s'. On the other hand, comparability requires only that the *relative gain* from being assigned to the selective instead of the non-selective school be the same for (x, s) and for some (x', s').

Figure 1 illustrates the difference between strong comparability and comparability in a simple example. It depicts a setting with 4 categories, x_1, x_2, x_3, x_4, and 2 positions, p_1, p_2, and plots expected surplus as a function of agents' signals. While this example is very stylized, the definitions of comparability allow for much more general and non-linear settings, where the mapping between signals for different categories is significantly less regular.

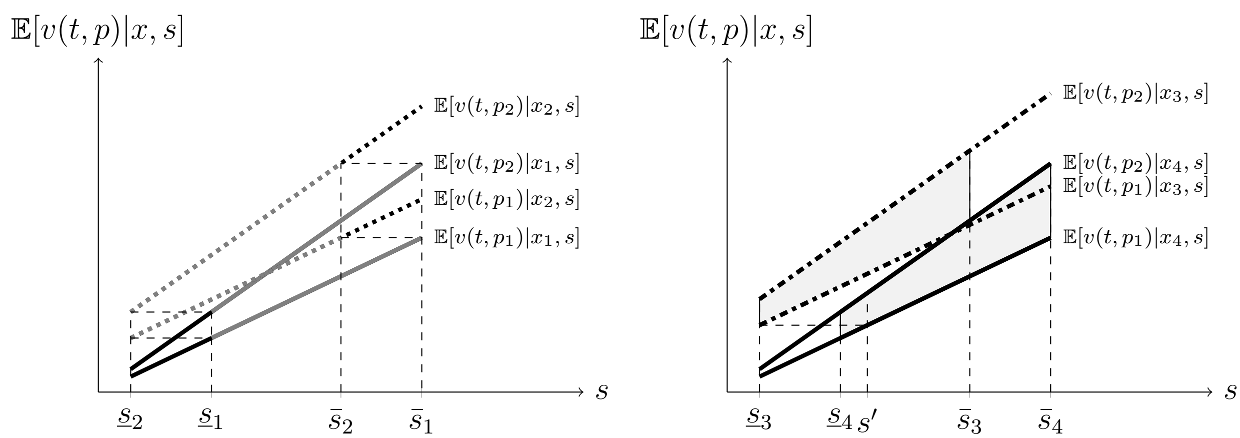

Figure 1: Strongly comparable and comparable categories. Left panel: x_1 and x_2 are strongly comparable. Right panel: x_3 and x_4 are comparable, but not strongly comparable.

In the left panel, categories x_1 (in solid lines) and x_2 (in dotted lines) are strongly comparable. To see this, note that for category x_1 and any signal $s \in [\underline{s}_1, \bar{s}_1] \subset S$ there exists some signal $s' \in [\underline{s}_2, \bar{s}_2] \subset S$ for category x_2 such that $\mathbb{E}[v(t,p)|x_1, s] = \mathbb{E}[v(t,p)|x_2, s']$ for both $p \in \{p_1, p_2\}$. The expected surpluses associated with these signals are highlighted in gray in the figure.[5] Strong comparability ensures that some signals are associated with the same absolute levels of expected surplus, regardless of which position p is used to make the comparison.

In the right panel, categories x_4 (in solid lines) and x_3 (in dash-dotted lines) are not strongly comparable, but are comparable. To see the former, note that there exist no signals s, s' such that the absolute levels of expected surplus are equal for both p_1 and p_2. That is, there are no s, s' such that the two solid lines have the same absolute levels as the two dash-dotted lines, when evaluated at s, s' respectively.[6] To see that comparability holds, consider the relative differences in expected surplus from p_1 and p_2; i.e. the difference between the two solid lines and the difference between the two dash-dotted lines. For category x_4 and any signal $s \in [\underline{s}_4, \bar{s}_4] \subset S$ there exists some signal $s' \in [\underline{s}_3, \bar{s}_3] \subset S$ for category x_3 such that $\mathbb{E}[v(t,p_2)|x_4,s] - \mathbb{E}[v(t,p_1)|x_4,s] = \mathbb{E}[v(t,p_2)|x_3,s'] - \mathbb{E}[v(t,p_1)|x_3,s']$. The differences in expected surpluses associated with these signals are highlighted in gray in the figure.[7]

A key feature of both comparability definitions above is that agents of different categories can be compared uniformly across all positions. With strong comparability, comparisons of (x, s) and (x', s') in terms of expected surplus must agree for all positions p. With comparability, comparisons of (x, s) and (x', s') in terms of relative gains in expected surplus must agree for all pairs of positions p, p', with $p > p'$. The next numerical example illustrates the difference between comparability and strong comparability in a different format, and also shows how comparability can fail to hold in a setting with more than 2 positions.

Example 1. *Consider 4 categories, x_1, x_2, x_3, x_4, and 3 positions, p_1, p_2, p_3. Each table below shows the expected surplus from the 3 positions (in columns) for one of the categories, for 3 possible signals (in rows), s_1, s_2, s_3.*

[5] Furthermore, for category x_1 and signals $s < \underline{s}_1$, case (iii) of the definition is satisfied. For category x_2 and signals $s > \bar{s}_2$, case (ii) of the definition is satisfied. Hence the definition of strong comparability is satisfied for all (x_1, s) and all (x_2, s).

[6] Take e.g. the signal $\underline{s}_3$: when considering position p_1, the signal $\underline{s}_3$ translates to some s' that yields the same expected surplus, but $\underline{s}_3$ and s' do not yield the same expected surplus when considering p_2. Specifically, we have $\mathbb{E}[v(t,p_1)|x_3,\underline{s}_3] = \mathbb{E}[v(t,p_1)|x_4,s']$ for some s', but $\mathbb{E}[v(t,p_2)|x_3,\underline{s}_3] \neq \mathbb{E}[v(t,p_2)|x_4,s']$. Note also that the expected surplus $\mathbb{E}[v(t,p_1)|x_3,\underline{s}_3]$ is strictly inside the range of $\mathbb{E}[v(t,p_1)|x_4,s]$, so the second and third cases of the definition of strong comparability also fail.

[7] Furthermore, for category x_4 and signals $s < \underline{s}_4$, case (iii) of the definition is satisfied. For category x_3 and signals $s > \bar{s}_3$, case (ii) of the definition is satisfied. Hence the definition of comparability is satisfied for all (x_1, s) and all (x_3, s).

Category x_1	p_1	p_2	p_3
$\mathbb{E}[v(t,p)\|x_1,s_1]$	*1*	*3*	*6*
$\mathbb{E}[v(t,p)\|x_1,s_2]$	*2*	*5*	*10*
$\mathbb{E}[v(t,p)\|x_1,s_3]$	*3*	*7*	*13*

Category x_2	p_1	p_2	p_3
$\mathbb{E}[v(t,p)\|x_2,s_1]$	*1*	*4*	*9*
$\mathbb{E}[v(t,p)\|x_2,s_2]$	*4*	*8*	*14*
$\mathbb{E}[v(t,p)\|x_2,s_3]$	*5*	*10*	*17*

Category x_3	p_1	p_2	p_3
$\mathbb{E}[v(t,p)\|x_3,s_1]$	*2*	*5*	*10*
$\mathbb{E}[v(t,p)\|x_3,s_2]$	*3*	*7*	*13*
$\mathbb{E}[v(t,p)\|x_3,s_3]$	*4*	*9*	*16*

Category x_4	p_1	p_2	p_3
$\mathbb{E}[v(t,p)\|x_4,s_1]$	*1*	*2*	*4*
$\mathbb{E}[v(t,p)\|x_4,s_2]$	*2*	*3.9*	*10*
$\mathbb{E}[v(t,p)\|x_4,s_3]$	*3*	*7*	*14*

Based on the signals and associated expected surplus depicted in the tables, categories x_1 and x_2 are comparable, x_1 and x_3 are strongly comparable, and x_1 and x_4 are not *comparable.*

First, categories x_1 and x_2 are comparable because

$$\mathbb{E}[v(t,p')|x_1,s_2] - \mathbb{E}[v(t,p)|x_1,s_2] = \mathbb{E}[v(t,p')|x_2,s_1] - \mathbb{E}[v(t,p)|x_2,s_1]\, \forall p,p',$$
$$\mathbb{E}[v(t,p')|x_1,s_3] - \mathbb{E}[v(t,p)|x_1,s_3] = \mathbb{E}[v(t,p')|x_2,s_2] - \mathbb{E}[v(t,p)|x_2,s_2]\, \forall p,p'.$$

I.e. (x_1,s_2) and (x_2,s_1) have the same relative gains from all pairs of positions, and (x_1,s_3) and (x_2,s_2) have the same relative gains. Moreover, (x_1,s_1) and (x_2,s_3) satisfy the conditions for comparability in cases (iii) and (ii) of the definition, respectively.

Second, categories x_1 and x_3 are strongly comparable because

$$\mathbb{E}[v(t,p)|x_1,s_2] = \mathbb{E}[v(t,p)|x_3,s_1]\, \forall p,$$
$$\mathbb{E}[v(t,p)|x_1,s_3] = \mathbb{E}[v(t,p)|x_3,s_2]\, \forall p.$$

I.e. (x_1,s_2) and (x_3,s_1) have the same expected surplus from all positions, and (x_1,s_3) and (x_2,s_2) have the same expected surplus. Moreover, (x_1,s_1) and (x_3,s_3) satisfy the conditions for strong comparability in cases (iii) and (ii) of the definition, respectively.

Finally, categories x_1 and x_4 are not comparable. To see this, note that

$$\mathbb{E}[v(t,p_2)|x_4,s_2] - \mathbb{E}[v(t,p_1)|x_4,s_2] < \mathbb{E}[v(t,p_2)|x_1,s] - \mathbb{E}[v(t,p_1)|x_1,s]\, \forall s,$$
$$\mathbb{E}[v(t,p_3)|x_4,s_2] - \mathbb{E}[v(t,p_2)|x_4,s_2] > \mathbb{E}[v(t,p_3)|x_1,s] - \mathbb{E}[v(t,p_2)|x_1,s]\, \forall s.$$

That is, using positions p_1 and p_2 to compare the relative gains in expected surplus, an agent with (x_4,s_2) has smaller *relative gains than any (x_1,s). On the other hand, using positions*

p_2 and p_3 to make the same comparison, that agent with (x_4, s_2) has larger *relative gains than any (x_1, s). Hence the comparison in terms of relative gains from different positions is not uniform across positions. Therefore the signal s_2 for category x_4 cannot be compared to any signal for category x_1. This counter-example for comparability requires that there be at least 3 positions or treatments. With binary assignments comparability always holds, as shown in section 3.3, although strong comparability need not hold.*

For the remainder of the analysis I will assume that all categories are comparable. I further study the two notions of comparability later in the paper. After presenting my main results in section 3.1 and several extensions and applications in section 3.2, in section 3.3 I provide three propositions which characterize different settings where comparability holds.

3.1 Optimal policy

Here I present the main results of the paper. Theorem 1 provides a general characterization of the optimal policy which maximizes expected total surplus. Theorem 2 shows when the optimal policy features differential treatment. Theorem 3 shows that differential treatment can be optimal because of differences in the informativeness or noisiness of signal structures across characteristics, not just because of biases in these signals.

Theorem 1. *Suppose Assumptions 1-3 hold and categories are comparable. The optimal policy $\mathscr{P}^*$ assigns agents to positions assortatively with respect to an index defined by $r(x,s) = \mathbb{E}[v(t,p_m)|x,s] - \mathbb{E}[v(t,p_{m-1})|x,s]$.*

In this theorem I first define an index function, given by $\mathbb{E}[v(t,p)|x,s] - \mathbb{E}[v(t,p')|x,s]$, for some $p > p'$, and show that this index induces a total preorder over $\{(x_i, s_i)\}$. I then show that comparability impies this preorder is invariant with respect to p and p'. Hence without loss of generality I let $p = p_m$ and $p' = p_{m-1}$. I then show that among policies that are deterministic (up to tie-breaking among index-equivalent agents), optimality implies that agents are assigned assortatively according to this index. Finally, I extend the result to stochastic policies, to show that there is no gain from randomization among non-payoff-equivalent agents; hence the optimal policy is essentially deterministic, and assigns agents to positions assortatively with respect to the index function, i.e. the optimal policy is weakly increasing in $r(x,s)$.

More generally, one can also find the optimal assignment in a setting without comparability, using the Hungarian assignment algorithm and recent developments from the engineering and computer science literatures that build on it. While such algorithms can solve for the optimal assignment in a less structured setting, they are less interpretable in terms of economic intuition and do not provide comparative statics, because they do not have an index implementation, unlike the result in Theorem 1. That is, the predictions of these algorithms are more of a black box.

Theorem 1 and the definition of comparability suggest an interesting avenue for further work: one can decentralize this assignment model so that a different principal is in charge of each position and decides who is admitted to it or not, in order to maximize the surplus generated in that position. Such a model is useful for studying decentralized college admissions, relative to the efficient assignment that is characterized in this paper. I discuss a decentralized version of the model in section 3.2.3. This extension considers a stylized setting that abstracts away from other possible frictions and issues that may arise with decentralization, but it provides a useful first step in this direction.

Prior work (e.g. Chan and Eyster, 2003) has shown that the optimal policy may in general be non-deterministic, including for reasons that are not driven by aggregate uncertainty. However the proof of Theorem 1 implies the following observation in the current setting.

Corollary 1. *The optimal policy is deterministic, except for tie-breaking among payoff-equivalent agents.*

Moreover, the optimal assignment described in Theorem 1 can be implemented with category-specific signal thresholds for each position. The proof of the result directly shows how to set these thresholds optimally.

Corollary 2. *For a set of agents with observables* $\{(x_i, s_i)\}$*, the optimal deterministic policy can be implemented with category-specific signal cutoffs for each position:* $\bar{s}_p(x) \equiv \min\{s_i : \mathbb{P}^*(p|x, s_i) > 0\}$*.*

Theorem 1 and Corollary 2 show that differential treatment arises endogenously in this environment, as the solution to the policy-maker's surplus maximization problem. In general, the category-specific cutoffs that implement $\mathscr{P}^*$ need not be the same across categories, and hence some categories of agents will be favored, relative to their signal realizations. In an education context, it may be efficient to admit students to schools or universities according to exam score cutoffs that are different across categories (e.g. across socio-economic status, race, or some other covariate). Similarly, in a labor market context, it may be efficient to hire or promote candidates according to performance thresholds that are different across categories (e.g. across gender or some other covariate).

Similar cutoff implementations are developed in other recent work on the design of education markets (e.g. Azevedo and Leshno, 2016; Agarwal and Somaini, 2018; Leshno and Lo, 2020; Leshno, 2021). However, the most important feature of this implementation is that cutoffs generally need to be able to vary within positions in order to implement an efficient assignment. Such variations explicitly account for differences in the informational content of signals for agents of different categories. For example, if two agents have characteristics and signals (x', s') and (x'', s'') such that $r(x', s') > r(x'', s'')$, then (x', s') will be assigned to a weakly higher position, regardless of whether $s' > s''$ or $s' < s''$.

Moreover, suppose that for categories x', x'' it is the case that $r(x', s) > r(x'', s)$ for all $s \in S$. The optimal assignment from Theorem 1 implies that an agent of category x' would always be assigned to a weakly higher position than an agent of category x'' who has the same signal, with strictly higher assignments for some sets of capacities $(k_1, ..., k_m)$. In other words, $p(x', s) \geq p(x'', s)$ when the optimal $p(x, s)$ is a singleton, or more generally $\text{supp}[\mathbb{P}^*(x', s)] \geq \text{supp}[\mathbb{P}^*(x'', s)]$ in the strong set order, and these supports overlap by at most one treatment. Such a pattern, where the benefit indices can be ordered across categories, emerges in some more structured environments where the conditional type distributions can be ordered across categories. As an application, in section 3.2.1 I consider a setting where the conditional type distributions can be ordered across categories in terms of first-order stochastic dominance, and show that categories with worse distributions are optimally assigned according to lower signal cutoffs for each position.

Theorem 1 and Corollary 2 imply that the optimal policy is assortative with respect to the agent's index, $r(x, s)$, but in general need not be assortative with respect to signals. The optimal assignment is monotonic in signals within categories, but not across categories. Note that assortativity with respect to signals would be equivalent to the optimal signal threshold for each position being category-invariant, i.e. $\bar{s}_p(x)$ would be constant in x for each p. Example 2 illustrates the non-monotonicity with respect to signals in a very simple setting.

Example 2. *Suppose there are 2 positions, p_1, p_2, with capacities $k_1 = 2$ and $k_2 = 1$, and 3* *agents with observables* $(x_1, s_1), (x_1, s_2), (x_2, s_1)$*, for some categories* x_1, x_2 *and some signals,* $s_1 < s_2$*, with corresponding expected values for the positions given by the following table:*

	p_1	p_2
$\mathbb{E}[v(t,p)\|x_1, s_1]$	1	2
$\mathbb{E}[v(t,p)\|x_1, s_2]$	2	4
$\mathbb{E}[v(t,p)\|x_2, s_1]$	0	3

The optimal assignment here is $p(x_2, s_1) = p_2$*,* $p(x_1, s_2) = p(x_1, s_1) = p_1$*. So* $p(x_2, s_1) > p(x_1, s_2)$ *even though* $s_1 < s_2$*, and despite the fact that* $\mathbb{E}[v(t, p_2)|x_2, s_1] < \mathbb{E}[v(t, p_2)|x_1, s_2]$*. Therefore the optimal policy is not increasing with respect to either the agents' signals or the agents' expected values.*

At this point it is natural to ask whether there are settings where the optimal assignment does *not* require differential treatment and can therefore be implemented with category-invariant cutoffs. Before presenting the answer to this question, I first provide a useful definition of when signal structures are equivalent or identical.[8] To state the definition, first let $\widetilde{v}_{p,x}(s) \equiv \mathbb{E}[v(t,p)|x,s]$ denote the expected value of $v(\cdot)$ with respect to $F(t|x,s)$, for any p and x. Then let $\widetilde{w}_{p,x}(v) \equiv \widetilde{v}^{-1}_{p,x}(s)$ denote the inverse, for any p and x. This inverse is well-defined, because $v(t,p)$ is increasing in t and the expectation is taken with distributions $F(t|x,s)$ which are FOSD-increasing in s.

Definition 3. *Two signal structures* $F_{t,x}$ *and* $F_{t,x'}$ *are* ***equivalent*** *with respect to* $v(\cdot)$ *if* $\exists \delta$ *s.t.*

$$\widetilde{w}_{p,x}(v) = \widetilde{w}_{p,x'}(v) + \delta$$

for any p*, for all* $v \in \mathrm{Im}[\widetilde{v}_{p,x}(s)] \cap \mathrm{Im}[\widetilde{v}_{p,x'}(s)]$*.*
The signals $F_{t,x}$ *and* $F_{t,x'}$ *are* ***identical*** *with respect to* $v(\cdot)$ *if this holds for* $\delta = 0$*.*

In words, two signal structures are *equivalent* if the expected values as a function of signals are horizontal translations of each other, over their common image. That is, the expected value $\widetilde{v}_{p,x}(s)$ corresponding to a signal distribution $F_{t,x}$ is a shift to the left or right of the expected value $\widetilde{v}_{p,x'}(s)$ corresponding to a signal distribution $F_{t,x'}$. These expected value functions only need to agree with each other over the range where their images intersect. Analogously, two signal structures are *identical* with respect to $v(\cdot)$ if their expected values

[8] In Proposition 7 I also show that this definition of equivalence is sufficient for strong comparability.

as a function of signals are exactly equal. This is the case if $F(t|x,s)$ is constant in x; i.e. t conditional on s is independent of x.

The next result answers the question of whether there are settings that do not require differential treatment.

Theorem 2. *Suppose Assumptions 1-3 hold and categories are comparable. The optimal policy* $\mathscr{P}^*$ *can be implemented with category-invariant cutoffs for any* $(k_1, ..., k_m)$ *if and only if the signals* $\{F_{t,x}\}$ *are identical with respect to* $v(\cdot)$*. In particular, if* $\{F_{t,x}\}$ *are equivalent but not identical,* $\exists\ x, x'$, $(k_1, ..., k_m)$ *and* $s > s'$ *such that* $\mathbb{P}^*(x,s) < \mathbb{P}^*(x',s')$.

I show that the optimal policy is assortative with respect to signals only in a very special case, when the signal structures are *identical* with respect to $v(\cdot)$. This is a very demanding assumption: it means that other characteristics represented by X have effectively no impact on the expected surplus that an agent generates as a function of their signal. In an education context this would mean, for example, that students with different socio-economic status or parental investment in schooling have the same expected returns to education. In a labor market context it would mean, for example, that factors such as parental leave, which are empirically known to affect men's and women's careers differently, have no impact on the expected surplus that an employee generates as a function of their past output or some other signal of productivity. More generally, one might only expect the optimal policy to be "quasi-assortative," in the sense that it assigns agents to positions according to their benefit indices, which correspond only loosely to their signals.

The result above illustrates a *bias* motive for differential treatment: to maximize surplus the policy-maker must translate possibly biased signals across categories. Since surplus is a function of types, not of signals per se, the policy-maker maps signals into beliefs about types, and this mapping may be different across categories.[9] The signal structures across different categories may be inherently biased in favor of some categories, unless these signal structures are identical, which is unlikely to be the case in practice. If the distributions that generate these signals are not identical, then the translation of signals into types will compensate agents from categories whose signals are inherently worse, giving rise to differential treatment, such as affirmative action.

[9] The definition of comparability ensures that such a mapping is possible—if categories are not comparable, then the optimal policy may assign agents to positions in a way that is not assortative with respect to any index across categories.

Example 3 below is an especially simple illustration of Definition 3, where the distributions of signals are horizontal translations of one another. But the equivalence can hold more generally—it would be sufficient, for example, that for every s there exists some $s' = s + \delta$ such that the Bayesian posteriors $F(t|x, s)$ and $F(t|x', s')$ are equal.

Example 3. *Suppose $t \in \{t_L, t_H\}$ and $S = [0, 1]$. Consider 3 categories, x_1, x_2, x_3 with signal distributions given by:*

$$F(s|t, x_1) = \begin{cases} \min\{3s, 1\} & \textit{for } t = t_L \\ \min\{(3s)^3, 1\} & \textit{for } t = t_H \end{cases}$$

$$F(s|t, x_2) = \begin{cases} \max\{0, \min\{3(s - \frac{1}{2}), 1\}\} & \textit{for } t = t_L \\ \max\{0, \min\{[3(s - \frac{1}{2})]^3, 1\}\} & \textit{for } t = t_H. \end{cases}$$

$$F(s|t, x_3) = \begin{cases} \max\{0, 3(s - \frac{2}{3})\} & \textit{for } t = t_L \\ \max\{0, [3(s - \frac{2}{3})]^3\} & \textit{for } t = t_H. \end{cases}$$

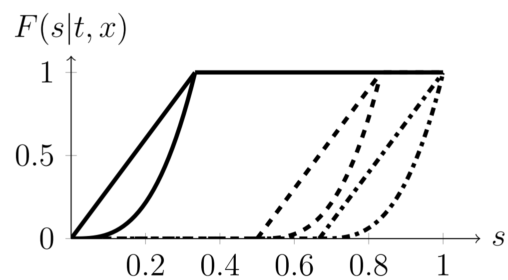

Figure 2: Signal distributions for x_1 (solid), x_2 (dashed) and x_3 (dash-dotted).

Note that F_{t,x_1}, F_{t,x_2} and F_{t,x_3} are pairwise equivalent, because a signal s for category x_1 corresponds to $s' = s + \frac{1}{2}$ for category x_2 and to $s'' = s + \frac{2}{3}$ for category x_3.

Consider a setting with 2 positions (p_1, p_2). Suppose first that agents belong to categories x_1 and x_3. A bias of size $\frac{2}{3}$ in the signal distributions means that the supports of agents' signals are disjoint here: $[0, \frac{1}{3}]$ for category x_1 and $[\frac{2}{3}, 1]$ for category x_3. Any agent of category x_3 would always have a higher signal than any agent of category x_1, although the same is generally not true of the benefit indices $r(x_3, s)$ and $r(x_1, s)$. In this case an optimal assignment would typically feature differential treatment for most capacities k_1 and k_2, because the signal bias is large by construction. I.e. an agent with (x_1, s) has the same conditional type distribution and the same expected surplus as an agent with $(x_3, s + \frac{2}{3})$.

On the other hand, suppose that agents belong to categories x_2 *and* x_3 *and have observables* $\{(x_2, 0.6), (x_2, 0.8), (x_3, 0.7), (x_3, 0.8)\}$*. Note that* $r(x_2, 0.8) > r(x_3, 0.8) > r(x_2, 0.6) > r(x_3, 0.7)$*. Here one can construct some examples that do not require differential treatment. For instance, if the capacity of* p_2 *is* $k_2 = 2$*, then it is optimal to assign* $(x_2, 0.8)$ *and* $(x_3, 0.8)$ *to* p_2 *and the rest of the agents to* p_1*. So the optimal assignment would be implemented with category-invariant signal cutoffs for position* p_2*:* $\bar{s}_{p_2}(x_2) = \bar{s}_{p_2}(x_3) = 0.8$*. Nonetheless, because there is a bias of size* $\frac{1}{6}$ *in the signal distributions for categories* x_2 *and* x_3*, there will be some capacities* (k_1, k_2) *such that the optimal assignment requires differential treatment. For example, if* $k_2 = 3$*, then it is optimal to assign* $(x_2, 0.8)$*,* $(x_3, 0.8)$ *and* $(x_2, 0.6)$ *to* p_2*, leaving* $(x_3, 0.7)$ *for position 1. That is, an agent with a signal of* 0.7 *is assigned to a lower position, while another agent with signal* 0.6 *is assigned to a higher one. The signal cutoffs for position* p_2 *obviously vary across categories:* $\bar{s}_{p_2}(x_2) = 0.6$*, while* $\bar{s}_{p_2}(x_3) = 0.8$*.*

The above example also illustrates Theorem 2. In this setting signals are biased, as the distributions of signals conditional on types are horizontal shifts of one another. The example shows that when the bias is small enough, one can construct examples where the optimal assignment does not require differential treatment for *some* capacities. However, it can also be the case that the optimal policy generically features differential treatment for any given capacities, if the bias is large enough. More generally, the only case where the optimal assignment does not require differential treatment regardless of capacities is that where the distributions $\{F_{t,x}\}$ are identical across x.

The next result highlights a separate *informativeness* motive for differential treatment: even if signals are unbiased across categories (i.e. a signal implies the same expected type for any category), the surplus-maximizing policy may favor agents whose categories have either more noisy or less noisy signals. This is because the optimal policy is assortative with respect to $r(x, s)$, which need not be constant in x even if $\mathbb{E}[t|x, s]$ is constant in x. In particular, if v exhibits some intrinsic preference for or against variability, then agents whose expected types are equal, conditional on the same signal, will be ranked differently according to $r(x, s)$.

This intuition is quite general and can be formalized in a more specific setting. I consider two classes of surplus functions which exhibit a preference for and against dispersion in types.

Definition 4. *A value function* $v(t, p)$ *has (strictly)* ***convex differences in*** t *if* $v(t, p') - v(t, p)$ *is a (strictly) convex function of* t*, for any* $p' > p$*. Alternatively,* $v(t, p)$ *has (strictly)* ***concave***

differences in *t if $v(t,p') - v(t,p)$ is a (strictly) concave function of t, for any $p' > p$.*[10]

Concave and convex differences in t represent the incremental value of assigning an agent to a higher position, as a function of that agent's type. Intuitively, convex differences in t mean the agent generates more surplus from being assigned to a higher position, and the extent of this is increasing in their type t. Such an assumption could represent settings where there are disproportionately large gains from "superstar" agents with very high types in high positions. In an education context, this would be the case if very high ability students generate larger-than-proportional values from higher positions, for example in the admission of future superstar scientists to postgraduate programs. In an organizational context, this would be the case if the ability or productivity of managers in higher-tier management positions has an increasingly larger effect on the organization's performance. Conversely, concave differences in t mean the agent generates more surplus from being assigned to a higher position, but the extent of this decreases in their type t. Such an assumption would capture diminishing returns to ability or productivity when comparing the incremental gains from being assigned to a higher position. This would mean that the policy-maker prefers consistency in the agents' types. Concave and convex differences in t are akin to risk-averse and risk-seeking preferences, and in my model the value function v determines whether the policy-maker has an intrinsic apreference for or against dispersion.

Next, I use these definitions to illustrate the informativeness motive for differential treatment, in a setting with unbiased signals with different informational content.

Theorem 3. *Suppose Assumptions 1-3 hold, categories are comparable, $\mathbb{E}[t|x,s]$ is constant in x for every s, and for some x', x'', $F(t|x'',s)$ is a mean-preserving spread of $F(t|x',s)$. If $v(t,p)$ has [strictly] convex differences in t, then $r(x'',s) \geq [>] r(x',s)$ for all s. If $v(t,p)$ has [strictly] concave differences in t, then $r(x'',s) \leq [<] r(x',s)$ for all s.*

Theorem 3 shows that the informational content of signals affects how agents are ranked even if they are unbiased. Because the benefit indices differ across categories, there exist problems where the capacities are such that the optimal policy assigns some agents with lower signals to higher positions, based on the differences in ranks across categories. That is, agents from categories with more or less noisy signals may have a higher or lower index,

[10]One can also define analogous properties of the differences when T is a discrete space, but for ease of exposition I will focus on the case where T is continuous.

depending on whether v has convex or concave differences in type. Noisiness is formalized in terms of mean preserving spreads in the conditional type distributions, which in this case is equivalent to the convex stochastic order and second-order stochastic dominance.

The literature on statistical discrimination has generally focused on special environments where groups with less noisy conditional type distributions optimally receive more favorable treatment. This is the case, for example, in Aigner and Cain (1977), who assume that 2 groups of employees have different variances of their signal distributions and an employer explicitly dislikes noise in the conditional type distribution; hence the group with the noisier signal is discriminated against. A similar prediction is made in Cornell and Welch (1996). Here prospective employees either have the same characteristic as the employer or a different one, and the employer is assumed to observe more signals about candidates with the same characteristic. By assumption, these candidates have a more dispersed distribution of signals in general, but the conditional distribution of their types given their signals is in fact more precise. Hence the employer optimally tends to discriminate in favor of the same-characteristic candidates with less noisy conditional type distributions. Similarly, Bjerk (2008) studies a model where two groups of job candidates have skill signals with different precision, and explicitly assumes a discrete form of concave differences. Hence the group with less noisy conditional type distributions is favored. However, these types of results in the previous literature can be reversed altogether. Theorem 3 means that differential treatment need not inherently disfavor agents with noisier conditional type distributions. This result shows that whether it does or does not depends precisely on whether v has concave or convex differences in t.

The following example illustrates the definition of concave and convex differences and the intuition from Theorem 3. I discuss further results in this direction in section 3.2.2, where I also consider aggregate inequality. In that section I also provide another parametric example with a university admissions interpretation.

Example 4. *Consider the problem of an employer who wants to assign two workers among two positions,* p_1, p_2*, each with a capacity of* 1*. Suppose the employer's value function is* $v(t, p_i) = A_i \cdot t^\alpha$*, where* A_i *is a coefficient associated with position* p_i *and* t *is a worker's type, with* $t \in [0, 1]$*. Note that categories are comparable because the positions are binary (as shown in Proposition 5 later). Furthermore, assume* $A_2 > A_1 > 0$ *and* $\alpha > 0$*, so that* v *is increasing in both arguments and supermodular in* (t, p)*. The incremental gain from assigning a worker of type* t *to* p_2 *instead of* p_1 *is* $v(t, p_2) - v(t, p_1) = (A_2 - A_1) \cdot t^\alpha$*. Clearly*

v has convex differences in t if $\alpha > 1$ and concave differences in t if $\alpha < 1$.

Suppose the two workers have characteristics x_1 and x_2 respectively, and assume that $F(t|x_2, s)$ is a mean-preserving spread of $F(t|x_1, s)$ for all s. Note that $r(x_1, s) > r(x_2, s)$ for all s if $\alpha < 1$, and conversely $r(x_1, s) < r(x_2, s)$ for all s if $\alpha > 1$. Define a cutoff $\bar{s}_2(s_1)$ such that $r(x_1, s_1) = r(x_2, \bar{s}_2(s_1))$. Ignoring ties, based on Theorem 1 the optimal policy assigns worker 1 to position p_2 if $s_2 < \bar{s}_2(s_1)$, or worker 2 to position p_2 if $s_2 > \bar{s}_2(s_1)$.

Recall that r is increasing in s, $r(x_1, s_1) = r(x_2, \bar{s}_2(s_1))$ by definition, and $r(x_1, s_1) > r(x_2, s_1)$ if $\alpha < 1$. So $\bar{s}_2(s_1) > s_1$ if $\alpha < 1$. In this case the optimal policy favors the category x_1 with less noisy signals, because it holds a worker of category x_2 to a higher bar for promotion to position p_2, since $\bar{s}_2(s_1) > s_1$ for all s_1.

Analogously, if $\alpha > 1$ then $r(x_1, s_1) < r(x_2, s_1)$ and therefore $\bar{s}_2(s_1) < s_1$. In this case the optimal policy favors the category x_2 with noisier signals, because it holds a worker of that category to a lower bar for promotion to position p_2, since $\bar{s}_2(s_1) < s_1$ for all s_1.

3.2 Applications and extensions

My main results are derived in a general and non-parametric setup. In the following subsections I consider some applications to settings with additional structure, which allow me to derive more specific results.

3.2.1 Differential treatment in university admissions with biased signals

In some environments signaling technologies may be systematically biased across categories. That is, for some categories the mapping between types and signals may be consistently higher in some sense, thus making the informational content of signals biased across categories. To model such an application, I assume that signal structures can also be ordered across categories (whereas in my main analysis I only assumed that signals can be compared according to the MLRP *within* categories). In particular, the distributions of types conditional on some signal s may be comparable across characteristics x in the FOSD sense.

In the context of university admissions, for example, it may be that the distribution of abilities, conditional on some university entrance score s, is FOSD-decreasing in a socio-economic status covariate x, such as household income or parental investment. That is, conditional on the same exam score, a lower socio-economic status student would have a higher distribution of ability if socio-economic status tends to raise students' exam scores in some sense. This may reflect the fact that, in order to achieve the same exam score, a lower socio-economic status student may have to overcome larger obstacles, requiring higher ability; or it may reflect the beneficial effect of parental investments in education on exam performance. The literature on human capital (see e.g. Todd and Wolpin, 2007; Heckman, 2011; Fiorini and Keane, 2014) documents that higher socio-economic status parents tend to contribute more time and money to their children's education, which would justify such an assumption. The assumption is also consistent with the idea of "belief flipping" in the theoretical models of Fryer (2007) and Bohren et al. (2019), and with the latter's experimental evidence of belief reversion. This setting can be formalized as follows.

Assumption 4. *Suppose X can be ordered in such a way that $F(\cdot|x', s) \succ_{FOSD} F(\cdot|x'', s)$ for $x' < x''$ and for all s.*

Assumption 4 says that characteristics can be ordered so that "higher" characteristics are more strongly associated with lower types, given the same signal. A sufficient condition for this assumption to hold would be that the family of distributions $\{F_{t,x}\}$ satisfies a version of the strict monotone likelihood ratio property in x,[11] which would imply Assumption 4, analogously to Proposition 1. This assumption can also easily represent settings where a student belongs to multiple categories, because the model allows X to be vector-valued. In this case x and x' are sets of characteristics, rather than a single-dimensional covariate.[12]

Assumption 4 could apply not just to characteristics such as ethnicity or gender, but more broadly to socio-economic characteristics such as family composition, household income, or geographic location. Some standardized tests, such as the SAT, attempt to be neutral with respect to race, but it is difficult to imagine that such tests are neutral with respect to other,

[11]Specifically: the family of signal distributions $\{F_{t,x}\}$ satisfies the strict inverse monotone likelihood ratio property in x if X can be ordered so that for $x'' > x'$, the ratio of the densities $\frac{f_{t,x'}(s)}{f_{t,x''}(s)}$ is increasing in t, for any s, whenever it is well-defined.

[12]It is important to note that the decision-maker in my model is sophisticated and aware of the biases and differential noisiness of signals for different categories, which is also typically the case in the statistical discrimination literature. This is a reasonable simplification in settings such as university admissions, where the decision-maker observes agents' outcomes ex post and can learn over time how the signals differ systematically across categories. In future work one can also consider the need for the decision-maker to learn about the heterogeneity of signals across characteristics.

non-protected socio-economic characteristics, such as household income, location, or family status. The fact that Harvard University ranks candidates using its own internal composite scores, which allegedly treat candidates with the same SAT scores differently across SES characteristics, does not necessarily mean that Harvard's policy is motivated by concerns about differential informativeness, but it is at least consistent with my theoretical results. Harvard's motivation for using a composite admission score is likely to be more complex and to include other concerns besides the informational content of standardized tests—such as diversity in the composition of its student body. Nonetheless, my results here suggest that the differential informational content of standardized tests across various SES covariates could be a factor in admissions, in addition to student diversity.

Following a similar logic to that in Theorem 1, Assumption 4 implies that the optimal policy features a form of differential treatment that favors lower categories.

Proposition 2. *Suppose Assumptions 1-4 hold and categories are comparable. The optimal policy* $\mathscr{P}^*$ *is implemented by category-specific cutoffs* $\bar{s}_p(x)$ *that are increasing in* x*.*

That is, conditional on the same signal s, a lower-category agent is assigned according to a lower cutoff for each position, compared to a higher-category agent, and hence is favored by the optimal assignment policy.[13] This is because the optimal policy assigns agents according to the ranking function $r(x, s)$, which is increasing in s and decreasing in x under the assumptions in this application of the model. Thus Proposition 2 implies that differential treatment in admissions maximizes surplus: lower-characteristics agents are admitted according to lower signal cutoffs for each position. This applies not only to race-based affirmative action, but also more broadly to differential treatment across other characteristics.

The above application to university admissions considers the optimal assignment policy of a single decision-maker. In decentralized university admissions, such as the system used in the United States, this model applies to the choices made by one university's admissions office (e.g. to admit or not admit applicants, or to assign scholarships or different levels of financial aid to applicants). This is the same approach as that used in some of the recent market design literature on college admissions, which models individual colleges' choice functions, rather

[13] This favoritism may not strictly benefit lower-category agents, if the realizations of observables $\{(x_i, s_i)\}$ and the capacities of the positions happen to be such that the differences in ranks are irrelevant near the cutoffs for each position. However, because $r(x, s)$ is strictly decreasing in x, there exist some capacities and realizations of the observables such that the optimal policy strictly favors lower-category agents.

than market-wide matching outcomes (see e.g. Echenique and Yenmez, 2015; Abdulkadiroğlu and Grigoryan, 2021). The model can also be applied to centralized university admissions systems, like those used in other countries (see e.g. Chen and Kesten, 2017; Hafalir et al., 2018). In section 3.2.3 I also consider efficient assignments in a decentralized setting, which can represent decentralized university admissions systems where multiple universities make their own admissions decisions.

3.2.2 Optimal treatment inequality with unbiased signals

My results so far have considered how a policy-maker should use the signals and characteristics of agents to implement an assignment that maximizes total surplus. In many empirically relevant settings I find that the optimal policy favors some groups, relative to their *signals*, in the sense that it treats signals differently across characteristics. But this need not imply that different groups receive different treatments relative to their *types*. For example, if signals for one group are mechanically shifted downwards, compared to those for another group, as in Example 3, it may that the mapping from types to positions induced by $\mathscr{P}^*$ is in fact the same across X—that is, no group is actually favored when one considers how types are treated, even though $\mathscr{P}^*$ treats groups differentially in terms of signals.

However, in some cases $\mathscr{P}^*$ does induce unequal treatment across groups, with respect to types, not just signals. Interestingly, this can also be the case when signals are unbiased across characteristics, i.e. $\mathbb{E}[t|x, s]$ is constant in x. In this section I study such a setting as an application of the model.

Assumption 5. *Suppose $\{F_{s,x}\}$ is such that $\mathbb{E}[t|x, s]$ is constant in x, for every s, and for some x', x'', $F(t|x'', s)$ is a mean-preserving spread of $F(t|x', s)$.*

Assumption 5 is the same as the setting of Theorem 3, and represents an environment where signals are unbiased across categories, but their noisiness differs across categories. Signals are noisier for category x'', in the sense that the distribution of types conditional on a signal s is a mean-preserving spread of the distribution of types for category x', conditional on the same signal. In an education context, this could be the case if university entrance scores are more or less noisy measures of ability for some categories than others; that is, conditional on the same score, one group may have higher or lower dispersion in ability.

Proposition 3. *Suppose Assumptions 1-3 and 5 hold, and categories are comparable. If $v(t,p)$ has convex [concave] differences in t, then $p^*(x'', s) \geq [\leq] p^*(x', s)$ for all s, with strict inequality for some capacities and observables. Moreover, for all p, $\{s : \mathbb{P}^*(p|x'', s) > 0\} \leq [\geq] \{s : \mathbb{P}^*(p|x', s) > 0\}$ in the strong set order.*

Proposition 3 shows that aggregate inequality of treatments across groups can be optimal even when signals are unbiased, with identical prior distributions of types across groups. This is driven by the fact that the optimal policy assigns agents according to their expected incremental gains from higher treatments, not according to their expected types. Therefore the treatments of agents from categories x' and x'' differ in aggregate, with x'' receiving higher treatments when $v(t,p)$ has convex differences in t, because x'' has noisier signals and the policy-maker has a preference for dispersion. As a result, for any given treatment p, agents from category x'' who are assigned to p have lower signals, leading to aggregate inequality, despite the prior distribution of types being the same across categories.

In the context of university admissions this leads to the conclusion that average exam scores or expected abilities need not be equal across groups for each treatment. Rather, an efficient admissions policy would equalize the expected incremental surplus from admission. Similarly, in the context of hiring and promotion policies in organizations, the result implies that past performance or expected productivity need not be equal across groups.

The optimal treatment inequality characterized in Proposition 3 can be illustrated with a parametric example.

Example 5. *Suppose a university is deciding which one of 2 students to admit, i.e. the positions are (p_1, p_2) with capacities $k_1 = k_2 = 1$. The students have normally distributed abilities, $t_i \sim N(0,1)$, and belong to categories x^1 and x^2. A category x^j student has a university admission score $s_i = t_i + \varepsilon_i^j$, where $\varepsilon_i^j \sim N(0, \sigma_j)$ is zero-mean, normally-distributed noise, with variance $\sigma_2 > \sigma_1$, for categories x^1, x^2. Note that $F(t_2|x^2, s_2 = s)$ is a mean-preserving spread of $F(t_1|x^1, s_1 = s)$, for any s.*

For any $v(t,p)$ that has strictly convex differences in t, conditional on $s_1 = s_2 = s$, the university would strictly prefer to admit student 2, since $r(x^2, s) > r(x^1, s)$, and for any s_2 in some neighborhood below s, the university would prefer to admit student 2. That is, there exists some threshold $\bar{s}_2(s_1) < s_1$ such that for a given s_1, the university prefers to admit student 2 if and only if $s_2 \geq \bar{s}_2(s_1)$.

Then student 2 is admitted with probability $\mathbb{P}(s_2 \geq \bar{s}_2(s_1)) = \int \mathbb{P}(s_2 - s_1 \geq \bar{s}_2(s_1) - s_1|s_1)dF(s_1)$*. Note that* $s_2 - s_1$ *is normally distributed, with mean* 0*, and for any* s_1*,* $\bar{s}_2(s_1) - s_1 < 0$*. Therefore the integrand is strictly larger than* $\frac{1}{2}$ *everywhere, and* $\mathbb{P}(s_2 \geq \bar{s}_2(s_1)) > \frac{1}{2}$*. That is, the optimal policy is more likely to admit student* 2 *than student* 1*. Moreover, given that the type distributions are the same ex ante, the expected ability of student* 2 *conditional on* 2 *being admitted is lower than the corresponding expected ability of student* 1 *conditional on* 1 *being admitted. Therefore aggregate inequality arises both in terms of admission scores, since* $\bar{s}_2(s_1) < s_1$*, and in terms of expected types who are admitted across different categories.*

3.2.3 Decentralized assignment with differential treatment

The model thus far considered the problem of optimal assignment from the perspective of a single policy-maker, such as a university deciding which applicants to admit, a school district assigning students across schools, or an organization deciding which candidates to hire or to promote within its hierarchy. Suppose instead that the assignment of agents to positions is decentralized, so that a different decision-maker decides which applicants to accept for each treatment or position.

To represent this case, I assume the following timing of the game: (1) each applicant applies costlessly to all positions; (2) each decision-maker $j \in \{1, ..., m\}$ decides which subset of agents, denoted A_j, to accept for position p_j, in order to maximize

$$\sum_{i \in A_j} \mathbb{E}[v(t_i, p_j | x_i, s_i)];$$

and (3) agents accept the highest position to which they are accepted.

When will this decentralized assignment game yield efficient outcomes, such as the optimal index assignment characterized in Theorem 1? It turns out that the answer relates closely to the notion of comparability defined in Definition 1.

Proposition 4. *Suppose Assumptions 1-3 hold and categories are comparable. The outcome of the decentralized assignment model is efficient if and only if categories are strongly comparable.*

This result illustrates the broader usefulness of the comparability concepts defined in this paper, and also demonstrates that the richer assignment framework of this model of differential treatment, with multiple treatments or positions, can be applied to study novel questions not explored in the existing statistical discrimination literature—such as whether decentralized assignment yields efficient outcomes, or how centralized and decentralized assignments compare. Such comparisons are useful for example in the context of student assignment within school districts, where the delegation of admissions decisions to individual schools (which have reputational concerns) results in inefficient outcomes if and only if student categories are not strongly comparable. Similarly, in the context of hiring and promotion within an organization, where different divisions or tiers of the organization may make their own hiring and promotion decisions, seeking to maximize the division's output, rather than the overall value generated within the organization, decentralized assignment results in inefficient outcomes if and only if worker categories are not strongly comparable.

Whether decentralized assignments are efficient or not is independent of whether such assignments feature differential treatment or not. It is possible that a decentralized assignment is both efficient and features differential treatment, e.g. if categories are strongly comparable but signals are not identical according to Definition 3. Or it could be efficient and not feature differential treatment, e.g. if signals are identical according to Definition 3. It is also possible that decentralized assignments are not efficient, but still feature differential treatment, e.g. if categories are not strongly comparable and signals are biased in a particular way. To see the latter, consider the setting depicted in the right panel of Figure 1. If assignments are decentralized, so that each position maximizes the expected surplus generated by its own agents, then any position would prefer an agent with characteristics x_3 and a signal $\bar{s}_3 - \delta$ for a small enough δ over another agent with characteristics x_4 and signal $\bar{s}_4$. For example, in a setting with 2 positions, each with a capacity of 1, and 2 agents with $(x_3, \bar{s}_3 - \delta)$ and $(x_4, \bar{s}_4)$, position p_2 would accept $(x_3, \bar{s}_3 - \delta)$. This outcome is not efficient (because the agent with $(x_4, \bar{s}_4)$ has marginally larger gains from p_2 relative to p_1), but it still features differential treatment, because $\bar{s}_3 - \delta < \bar{s}_4$. Exploring differential treatment in a decentralized assignment framework could therefore be a fruitful direction for future research, independent of its efficiency implications.

3.3 Conditions for comparability

The concept of comparability is a very general one and holds in many environments—for example, most parametric models of statistical discrimination satisfy it. In this subsection

I provide three propositions which characterize multiple natural settings where it holds: (i) when the set of positions is binary; (ii) when $v(\cdot)$ satisfies an additional condition, which ensures that categories are comparable for any $\{F_{t,x}\}$; and (iii) when $\{F_{t,x}\}$ satisfies Definition 3, which ensures that categories are comparable for any $v(\cdot)$.

Proposition 5. *Suppose Assumptions 1-3 hold and positions are binary, $\{p_1, p_2\}$. Then all categories are comparable.*

With a binary set of positions the assumptions made earlier ensure that categories are comparable. These assumptions ensure that $\mathbb{E}[v(t,p)|x,s]$ is increasing in s for any p, and moreover $\mathbb{E}[v(t,p_2)|x,s] - \mathbb{E}[v(t,p_1)|x,s]$ is also continuous and increasing in s. Therefore a violation of comparability requires at least 3 positions among which agents must be assigned.

There are several applications where the set of treatments that the policy-maker must assign can be modeled as binary: the decision to admit or not admit a student in school and university admissions, as well as hiring decisions in organizations. In such settings, categories are comparable without any further assumptions on $v(t,p)$ or $\{F_{t,x}\}$.

The next proposition provides another environment where comparability holds.

Proposition 6. *Suppose Assumptions 1-3 hold and $v(t,p)$ is multiplicatively separable in t and p. Then all categories are strongly comparable, for any family of signals $\{F_{t,x}\}$.*

Proposition 6 shows that comparability is satisfied for *any* family of distributions $\{F_{t,x}\}$, as long as $v(\cdot)$ is multiplicatively separable. This makes comparability a widely applicable concept. Such separable value functions are commonly used in auction theory, mechanism design, and other settings, as they provide easy ways to model complementarities and supermodularity. For example, the value function in a quasi-linear environment is $v \cdot x$, where v is the agent's type and x is the allocation; the equivalent here is $v(t,p) = t \cdot p$.

The notion of signal equivalence in Definition 3 is also useful because it provides another type of sufficient condition for categories to be comparable, without imposing additional conditions on $v(\cdot)$.

Proposition 7. *Suppose Assumptions 1-3 hold and the signal distributions $\{F_{t,x}\}$ are (pairwise) equivalent, as in Definition 3. Then all categories are strongly comparable.*

Proposition 7 establishes a close connection between the equivalence of signals and the comparability of categories. Intuitively, if a signal s, for given category x, is equivalent to a signal s', for given category x', then the two categories are comparable.

4 Discussion

I study a setting where a policy-maker wants to assign agents with unobservable types and observable signals and characteristics to different positions or treatments. I characterize the policy that maximizes expected total surplus: it assigns agents to positions monotonically with respect to an index function that measures the expected incremental gains from different treatments. However, the optimal policy is in general not monotonic with respect to the agents' signals, or even the agents' expected types. Therefore the optimal policy will feature differential treatment in a variety of cases, where the distributions of signals conditional on types vary across characteristics.

I highlight two main intuitions for these results. First, to maximize surplus the policy-maker must translate signals into beliefs about types, and this translation need not be the same across characteristics if some groups of agents have different signal distributions—i.e. if signals are biased across characteristics. In this case the policy-maker may optimally favor groups whose signal distributions are worse. Second, differential treatment can be optimal even if signals are unbiased across characteristics, provided their informativeness differs across characteristics. That is, different groups may have more or less noisy signal distributions, and the optimal policy may favor some groups depending on whether the policy-maker's objective function has a preference for or against dispersion in types. These results provide a novel efficiency-based rationale for differential treatment policies.

The paper highlights several observations for public policy debates and empirical research on affirmative action and gender equity in education and labor market settings. First, differential treatment may arise out of concern for economic efficiency, not just fairness. Public discourse on affirmative action and on gender equity policies, for example, often focuses on fairness.[14] Second, comparing signals such as university exam scores or past

[14]This is not to suggest that there must exist a trade-off between fairness and efficiency; indeed, Persico (2002) implies that this need not be the case. The literature on "tagging" (Akerlof, 1978) also provides examples of other settings where differential treatment can improve efficiency.

employee performance across characteristics may be asking the wrong question, as it focuses on the wrong metric. In the context of maximizing surplus, it is expected incremental benefits that matter, not signals per se. Therefore a more relevant empirical question would be to estimate the returns to education across different characteristics for the marginal admitted students at a university, or the productivity across different characteristics for the marginal employees at an organization. If these are consistent across characteristics, this would point to an efficiency rationale for differential treatment. Third, comparing expected abilities or productivities across characteristics, conditional on exam scores or past performance, may also be asking the wrong questions, as the optimal policy can depend on the dispersion of types across characteristics, not just on means. So a more relevant empirical question would be to estimate whether there is inherently a disproportional benefit or loss due to dispersion in the abilities of students or in the productivity of employees. If so, this would also point to an efficiency rationale for differential treatment.

The paper contributes to the literature on matching and assignment games. The model explicitly distinguishes between an agent's signal and their payoff-relevant type, unlike most prior work in this area, which treats signals as directly payoff-relevant. The theoretical literature on college admissions, for example, typically models a university's payoffs over students' exam scores, rather than over some more primitive and welfare-relevant type, such as the returns to education. Similarly, in school choice papers, schools' priorities are sometimes defined based on students' exam scores. Distinguishing between signals and types is especially important for efficiency considerations, because the mapping between these need not be independent of other demographic or socio-economic characteristics.

The paper also contributes to the statistical discrimination literature in several ways. First, I focus on the problem of designing assignment policies which maximize efficiency, rather than on explaining discrimination per se, which is the main aim of statistical discrimination theories that focus on profit maximization rather than social surplus. For example, the literature on labor discrimination seeks to explain why an employer may favor one group of workers over another, whereas I consider how an employer who understands that some groups have less favorable signaling technologies may optimally treat workers in a way that compensates for such differences.

Second, I study a more general assignment model, with arbitrary treatments or positions, without making any specific parametric assumptions about the distributions of types and signals or the functional form of payoffs, in contrast to the existing literature on discrimination.

I provide general results using tools from information economics, monotone comparative statics, and stochastic ordering theory. This has several advantages relative to previous results. The analysis does not rely on any parametric or functional form assumptions, which models typically rely on for tractability, and this framework can also be analyzed without the assumption of comparability, which is implicitly guaranteed in existing models. The general treatment structure of this assignment model applies to a broader class of settings than the binary model usually considered in the literature—e.g. the assignment of students to schools within school districts and the promotion of employees to different tiers of an organizational hierarchy. My results on the effect of signal informativeness in section 3.2.2 show that both more and less informative signal structures can lead to certain categories of agents being favored, whereas the existing literature has focused on particular parametric assumptions which yield one direction of this relationship.

Finally, I show that this more general assignment framework enables the study of new settings not considered in the existing literature. For instance, I study the differences between centralized and decentralized assignments and the efficiency implications of decentralization. This framework also enables future work on differential treatment with endogenous human capital, on the effect of organizational structure on the extent of differential treatment, and on wage inequality in organizations.

5 References

Abdulkadiroğlu, Atila (2005) "College admissions with affirmative action," *International Journal of Game Theory*, Vol. 33, No. 4, pp. 535–549.

Abdulkadiroğlu, Atila and Aram Grigoryan (2021) "Priority-based Assignment with Reserves and Quotas," *Working Paper*.

Abdulkadiroğlu, Atila and Tayfun Sŏnmez (2003) "School Choice: A Mechanism Design Approach," *American Economic Review*, Vol. 93, No. 3, pp. 729–747.

Agarwal, Nikhil and Paulo Somaini (2018) "Demand Analysis Using Strategic Reports: An Application to a School Choice Mechanism," *Econometrica*, Vol. 86, No. 2, pp. 391–444.

Aigner, Dennis J. and Glen G. Cain (1977) "Statistical Theories of Discrimination in Labor Markets," *ILR Review*, Vol. 30, No. 2, pp. 175–187.

Akerlof, George A (1978) "The economics of "tagging" as applied to the optimal income tax, welfare programs, and manpower planning," *American Economic Review*, Vol. 68, No. 1, pp. 8–19.

Alesina, Alberto, Johann Harnoss, and Hillel Rapoport (2016) "Birthplace diversity and economic prosperity," *Journal of Economic Growth*, Vol. 21, No. 2, pp. 101–138.

Arrow, Kenneth J. (1973) *The Theory of Discrimination*, pp. 3–33: Princeton University Press.

Azevedo, Eduardo M. and Jacob D. Leshno (2016) "A Supply and Demand Framework for Two-Sided Matching Markets," *Journal of Political Economy*, Vol. 124, No. 5, pp. 1235–1268.

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

Besley, Timothy and Ravi Kanbur (1988) "Food subsidies and poverty alleviation," *The Economic Journal*, Vol. 98, No. 392, pp. 701–719.

Bjerk, David (2008) "Glass ceilings or sticky floors? Statistical discrimination in a dynamic model of hiring and promotion," *The Economic Journal*, Vol. 118, No. 530, pp. 961–982.

Blau, Francine D. and Lawrence M. Kahn (2017) "The Gender Wage Gap: Extent, Trends, and Explanations," *Journal of Economic Literature*, Vol. 55, No. 3, pp. 789–865.

Bodoh-Creed, Aaron L. and Brent R. Hickman (2018) "College assignment as a large contest," *Journal of Economic Theory*, Vol. 175, pp. 88–126.

Bohren, J. Aislinn, Alex Imas, and Michael Rosenberg (2019) "The Dynamics of Discrimination: Theory and Evidence," *American Economic Review*, Vol. 109, No. 10, pp. 3395–3436.

Chambers, Christopher P and Federico Echenique (2021) "A Characterisation of 'Phelpsian' Statistical Discrimination," *The Economic Journal*, Vol. 131, No. 637, pp. 2018–2032.

Chan, Jimmy and Erik Eyster (2003) "Does banning affirmative action lower college student quality?" *American Economic Review*, Vol. 93, No. 3, pp. 858–872.

Che, Yeon-Koo and Youngwoo Koh (2016) "Decentralized College Admissions," *Journal of Political Economy*, Vol. 124, No. 5, pp. 1295–1338.

Chen, Yan and Onur Kesten (2017) "Chinese College Admissions and School Choice Reforms: A Theoretical Analysis," *Journal of Political Economy*, Vol. 125, No. 1, pp. 99–139.

Chung, Kim-Sau (2000) "Role models and arguments for affirmative action," *American Economic Review*, Vol. 90, No. 3, pp. 640–648.

Coady, David, Margaret E Grosh, and John Hoddinott (2004) *Targeting of transfers in developing countries: Review of lessons and experience*: World Bank Publications.

Coate, Stephen and Glenn C Loury (1993) "Will affirmative-action policies eliminate negative stereotypes?" *American Economic Review*, pp. 1220–1240.

Cornell, Bradford and Ivo Welch (1996) "Culture, information, and screening discrimination," *Journal of Political Economy*, Vol. 104, No. 3, pp. 542–571.

Echenique, Federico and M. Bumin Yenmez (2015) "How to Control Controlled School Choice," *American Economic Review*, Vol. 105, No. 8, pp. 2679–94.

Ehlers, Lars, Isa E. Hafalir, M. Bumin Yenmez, and Muhammed A. Yildirim (2014) "School choice with controlled choice constraints: Hard bounds versus soft bounds," *Journal of Economic Theory*, Vol. 153, pp. 648 – 683.

Epple, Dennis, Richard Romano, and Holger Sieg (2008) "Diversity and affirmative action in higher education," *Journal of Public Economic Theory*, Vol. 10, No. 4, pp. 475–501.

Fang, Hanming and Andrea Moro (2011) "Theories of statistical discrimination and affirmative action: A survey," in *Handbook of Social Economics*, Vol. 1: Elsevier, pp. 133–200.

Fang, Hanming and Peter Norman (2006) "Government-Mandated Discriminatory Policies: Theory and Evidence," *International Economic Review*, Vol. 47, No. 2, pp. 361–389.

Fiorini, Mario and Michael P Keane (2014) "How the allocation of children's time affects cognitive and noncognitive development," *Journal of Labor Economics*, Vol. 32, No. 4, pp. 787–836.

Fryer, Roland G (2007) "Belief flipping in a dynamic model of statistical discrimination," *Journal of Public Economics*, Vol. 91, No. 5-6, pp. 1151–1166.

Fryer, Roland G and Glenn C Loury (2007) "Valuing identity: The simple economics of affirmative action policies."

Fu, Qiang (2007) "A Theory of Affirmative Action in College Admissions," *Economic Inquiry*, Vol. 44, No. 3, pp. 420–428.

Garg, Nikhil, Hannah Li, and Faidra Monachou (2021) "Dropping Standardized Testing for Admissions: Differential Variance and Access," *Working Paper*.

Hafalir, Isa E., Rustamdjan Hakimov, Dorothea Kübler, and Morimitsu Kurino (2018) "College admissions with entrance exams: Centralized versus decentralized," *Journal of Economic Theory*, Vol. 176, pp. 886–934.

Hafalir, Isa E., Fuhito Kojima, and M. Bumin Yenmez (2019) "Interdistrict School Choice: A Theory Of Student Assignment," *Working Paper*.

Hafalir, Isa E, M Bumin Yenmez, and Muhammed A Yildirim (2013) "Effective affirmative action in school choice," *Theoretical Economics*, Vol. 8, No. 2, pp. 325–363.

Hansen, Zeynep, Hideo Owan, and Jie Pan (2015) "The impact of group diversity on class performance: evidence from college classrooms," *Education Economics*, Vol. 23, No. 2, pp. 238–258.

Hartocollis, Anemona (2018) "Asian-Americans Suing Harvard Say Admissions Files Show

Discrimination," *Supreme Court Preview*, pp. 552–554.

Heckman, James J (2011) "The economics of inequality: The value of early childhood education.," *American Educator*, Vol. 35, No. 1, p. 31.

Herring, Cedric (2009) "Does diversity pay?: Race, gender, and the business case for diversity," *American Sociological Review*, Vol. 74, No. 2, pp. 208–224.

Hoogendoorn, Sander, Hessel Oosterbeek, and Mirjam van Praag (2013) "The Impact of Gender Diversity on the Performance of Business Teams: Evidence from a Field Experiment," *Management Science*, Vol. 59, No. 7, pp. 1514–1528.

Kojima, Fuhito (2012) "School choice: Impossibilities for affirmative action," *Games and Economic Behavior*, Vol. 75, No. 2, pp. 685 – 693.

Leshno, Jacob D (2021) "Stable Matching with Peer-Dependent Preferences in Large Markets: Existence and Cutoff Characterization," *Working Paper*.

Leshno, Jacob D and Irene Lo (2020) "The Cutoff Structure of Top Trading Cycles in School Choice," *The Review of Economic Studies*, Vol. 88, No. 4, pp. 1582–1623.

Levine, Sheen S., Evan P. Apfelbaum, Mark Bernard, Valerie L. Bartelt, Edward J. Zajac, and David Stark (2014) "Ethnic diversity deflates price bubbles," *Proceedings of the National Academy of Sciences*, Vol. 111, No. 52, pp. 18524–18529.

Lundberg, Shelly J and Richard Startz (1983) "Private discrimination and social intervention in competitive labor market," *American Economic Review*, Vol. 73, No. 3, pp. 340–347.

Mailath, George J, Larry Samuelson, and Avner Shaked (2000) "Endogenous inequality in integrated labor markets with two-sided search," *American Economic Review*, Vol. 90, No. 1, pp. 46–72.

Matysiak, Anna and Ewa Cukrowska-Torzewska (2021) "Gender and labour market outcomes," in *Research Handbook on the Sociology of the Family*: Edward Elgar Publishing.

Milgrom, Paul R (1981) "Good news and bad news: Representation theorems and applications," *The Bell Journal of Economics*, pp. 380–391.

Moro, Andrea and Peter Norman (2003) "Empirical implications of statistical discrimination on the returns to measures of skills," *Annales d'Economie et de Statistique*, pp. 399–417.

——— (2004) "A general equilibrium model of statistical discrimination," *Journal of Economic Theory*, Vol. 114, No. 1, pp. 1–30.

Nathan, Max and Neil Lee (2013) "Cultural Diversity, Innovation, and Entrepreneurship: Firm-level Evidence from London," *Economic Geography*, Vol. 89, No. 4, pp. 367–394.

Norman, Peter (2003) "Statistical discrimination and efficiency," *The Review of Economic Studies*, Vol. 70, No. 3, pp. 615–627.

Olszewski, Wojciech and Ron Siegel (2019) “Pareto Improvements in the Contest for College Admissions,” *Working Paper*.

Persico, Nicola (2002) “Racial Profiling, Fairness, and Effectiveness of Policing,” *American Economic Review*, Vol. 92, No. 5, pp. 1472–1497.

Phelps, Edmund S (1972) “The statistical theory of racism and sexism,” *American Economic Review*, pp. 659–661.

Todd, Petra E and Kenneth I Wolpin (2007) “The production of cognitive achievement in children: Home, school, and racial test score gaps,” *Journal of Human Capital*, Vol. 1, No. 1, pp. 91–136.

A Proofs

Theorem 1

Proof. Consider any positions p, p' with $p > p'$ and let

$$r_{p,p'}(x,s) = \mathbb{E}[v(t,p)|x,s] - \mathbb{E}[v(t,p')|x,s]$$

be the index of an agent with observables (x, s).

Note that $r_{p,p'}(\cdot)$ induces a total preorder on $X \times S$, with totally ordered equivalence classes, denoted by C_l. Then (x_i, s_i) and (x_j, s_j) belong to the same equivalence class iff $r_{p,p'}(x_i, s_i) = r_{p,p'}(x_j, s_j)$, and $(x_i, s_i) \in C_l$ and $(x_j, s_j) \in C_{l'}$ with $l < l'$ iff $r_{p,p'}(x_i, s_i) < r_{p,p'}(x_j, s_j)$.

Furthermore, for any positions $\hat{p}$ and $\hat{p}'$ with $\hat{p} > \hat{p}'$, let $r_{\hat{p},\hat{p}'}(x,s) = \mathbb{E}[v(t,\hat{p})|x,s] - \mathbb{E}[v(t,\hat{p}')|x,s]$ be an alternative index, which also induces a total preorder on $X \times S$. Because all categories x and x' are comparable, by definition we have that $r_{p,p'}(x_i, s_i) \geq r_{p,p'}(x_j, s_j)$ if and only if $r_{\hat{p},\hat{p}'}(x_i, s_i) \geq r_{\hat{p},\hat{p}'}(x_j, s_j)$. Hence the equivalence classes defined by $r_{p,p'}(\cdot)$ and $r_{\hat{p},\hat{p}'}(\cdot)$ are the same, and total preorders induced by $\mathbb{E}[v(t,p)|x,s] - \mathbb{E}[v(t,p')|x,s]$ are invariant with respect to (p, p'). So without loss of generality let the index be defined by

$$r(x,s) = \mathbb{E}[v(t,p_m)|x,s] - \mathbb{E}[v(t,p_{m-1})|x,s].$$

First, I show that in any deterministic policy $\mathscr{P}$, if agents are not assigned to policies in weakly increasing order of their index, the policy is suboptimal.

Consider any (x_i, s_i) and (x_j, s_j) with $r(x_i, s_i) < r(x_j, s_j)$. Suppose a policy $\mathscr{P}$ assigns $\mathbb{P}(x_j, s_j) = p'$ and $\mathbb{P}(x_i, s_i) = p''$ for some $p' < p''$.

Since $r(x_i, s_i) < r(x_j, s_j)$, we have $\mathbb{E}[v(t, p'')|x_i, s_i] - \mathbb{E}[v(t, p')|x_i, s_i] < \mathbb{E}[v(t, p'')|x_j, s_j] - \mathbb{E}[v(t, p')|x_j, s_j]$. Switching the assignment of agents i and j, holding all other assignments fixed, produces another feasible policy, with strictly higher total surplus, since

$$\mathbb{E}[v(t, p'')|x_i, s_i] + \mathbb{E}[v(t, p')|x_j, s_j] < \mathbb{E}[v(t, p'')|x_j, s_j] + \mathbb{E}[v(t, p')|x_i, s_i],$$

where the left-hand side is the i and j surplus under $\mathscr{P}$, and the right-hand side is the i and j surplus under the alternative policy.

Second, I show that there is no gain from randomization of non-index-equivalent agents within each position.

Consider any policy that randomizes among non-equivalent agents. Take the largest position p'' such that $q_i \equiv \mathbb{P}[p(x_i, s_i) = p''] \in (0, 1)$ and $q_j \equiv \mathbb{P}[p(x_j, s_j) = p''] \in (0, 1)$, for some agents (x_i, s_i) and (x_j, s_j) with $r(x_i, s_i) < r(x_j, s_j)$. Optimality requires that agents are matched to a position with probability 1, so $\exists$ another position $p' < p''$ s.t. $q'_j \equiv \mathbb{P}[p(x_j, s_j) = p'] \in (0, 1)$. Let $\hat{q} = \min\{q_i, q_j, q'_j\}$, and consider an alternative policy that moves a probability mass $\hat{q}$ of agent j's assignment from position p' to position p'', and a probability mass $\hat{q}$ of agent i's assignment from position p'' to position p'. Such a policy is feasible, holding all else constant. Because $r(x_i, s_i) < r(x_j, s_j)$, we have $\mathbb{E}[v(t, p'')|x_i, s_i] - \mathbb{E}[v(t, p')|x_i, s_i] < \mathbb{E}[v(t, p'')|x_j, s_j] - \mathbb{E}[v(t, p')|x_j, s_j]$. Hence the alternative policy yields strictly larger expected total surplus, since

$$\hat{q}\,\mathbb{E}[v(t, p'')|x_i, s_i] + \hat{q}\,\mathbb{E}[v(t, p')|x_j, s_j] < \hat{q}\,\mathbb{E}[v(t, p'')|x_j, s_j] + \hat{q}\,\mathbb{E}[v(t, p')|x_i, s_i],$$

where the left-hand side is the relevant i and j surplus under $\mathscr{P}$, and the right-hand side is the relevant i and j surplus under the alternative policy. Therefore the optimal policy cannot randomize among agents with different indices who are assigned to the same position.

Hence the optimal policy is weakly increasing in $r(x, s)$ and assigns agents assortatively with respect to $r(x, s)$, up to randomization among r-equivalent agents. □

Theorem 2

Proof. Proving sufficiency is straight-forward: if the signals $\{F_{t,x}\}$ are identical with respect to $v(\cdot)$, then for any $(k_1, ..., k_m)$ the optimal policy $\mathscr{P}^*$ can be implemented with category-invariant cutoffs.

To see this, suppose $\{F_{t,x}\}$ are identical with respect to $v(\cdot)$. Then $\widetilde{w}_{p,x}(v) = \widetilde{w}_{p,x'}(v)$ for any x, x', hence $\widetilde{v}_{p,x}(s) = \widetilde{v}_{p,x'}(s)$ for any x, x'. Therefore $r(x,s) = r(x',s)$ for any x, x'. Hence the optimal policy $\mathscr{P}^*$ has $\mathbb{P}^*(x,s) = \mathbb{P}^*(x',s)$, and hence $\bar{s}_p(x) = \bar{s}_p(x')$.

Next, I prove necessity by contrapositive: if the signals $\{F_{t,x}\}$ are not identical, then there exist $(k_1, ..., k_m)$ such that the optimal policy $\mathscr{P}^*$ is implemented with cutoffs that vary across categories.

Suppose $\{F_{t,x}\}$ are not identical. Then for some x', x'' and p, $\exists v^* \in \mathrm{Im}[\widetilde{v}_{p,x'}(s)] \cap \mathrm{Im}[\widetilde{v}_{p,x''}(s)]$ such that $\widetilde{w}_{p,x'}(v^*) \neq \widetilde{w}_{p,x''}(v^*)$. Note that $\widetilde{v}_{p,x}(s)$ is continuous, so there exists a neighborhood around $\widetilde{w}_{p,x'}(v^*)$ and around $\widetilde{w}_{p,x''}(v^*)$ such that $\widetilde{w}_{p,x'}(v) \neq \widetilde{w}_{p,x''}(v)$ in the neighborhood. Consider the largest such neighborhood of signals, denoted (s_{min}, s_{max}), where $\widetilde{v}_{p,x'}(s) \neq \widetilde{v}_{p,x''}(s)$ for any $s \in (s_{min}, s_{max})$. W.l.o.g. suppose $\widetilde{v}_{p,x'}(s) > \widetilde{v}_{p,x''}(s)$ and $r(x', s) > r(x'', s)$ for all $s \in (s_{min}, s_{max})$.

Let $k_m = 1 - F_{t,x'}(s_{min}) + \sum_{x \neq x'}[1 - F_{t,x}(s_{max})]$.

Then by Theorem 1, for $s \in (s_{min}, s_{max})$ the optimal policy $\mathscr{P}^*$ assigns $p^*(x', s) = p_m$, while $p^*(x'', s) < p_m$. Hence $\mathscr{P}^*$ is implemented by cutoffs $\bar{s}_{p_m}(x') < \bar{s}_{p_m}(x'')$, i.e. the optimal cutoffs vary across categories.

Finally, the above implies that in particular if $\{F_{t,x}\}$ are equivalent, but not identical, then there exist capacities such that the optimal cutoffs vary across categories. □

Theorem 3

Proof. Suppose $\{F_{s,x}\}$ is such that $\mathbb{E}[t|x,s]$ is constant in x, for every s, and for some x', x'', $F(t|x'', s)$ is a mean-preserving spread of $F(t|x', s)$.

Suppose $v(t, p)$ has convex [strictly convex] differences in t. Then

$$\begin{aligned} r(x'', s) &= \mathbb{E}[v(t, p_m)|x'', s] - \mathbb{E}[v(t, p_{m-1})|x'', s] \\ &= \int_t v(t, p_m) - v(t, p_{m-1}) dF(t|x'', s) \\ &\geq [>] \int_t v(t, p_m) - v(t, p_{m-1}) dF(t|x', s) \\ &= r(x', s) \end{aligned}$$

where the [strict] inequality follows from the fact that $v(t, p_m) - v(t, p_{m-1})$ is a convex function of t and $F(t|x'', s)$ is a mean-preserving spread of $F(t|x', s)$.

The proof that $r(x'', s) \leq [<] r(x', s)$ if $v(t, p)$ has concave [strictly concave] differences in t is analogous. □

Proposition 2

Proof. First, by Theorem 1, the optimal policy $\mathscr{P}^*$ is assortitative with respect to the index $r(x, s) \equiv \mathbb{E}[v(t, p_m)|x, s] - \mathbb{E}[v(t, p_{m-1})|x, s]$.

Second, $r(x, s) = \int_T [v(t, p_m) - v(t, p_{m-1})] dF(t|x, s)$ is increasing in s, because $v(t, p_m) - v(t, p_{m-1})$ is increasing in t, and under Assumption 3 $F(t|x, s)$ is FOSD-increasing in s.

Third, $r(x, s) = \int_T [v(t, p_m) - v(t, p_{m-1})] dF(t|x, s)$ is decreasing in x, because $v(t, p_m) - v(t, p_{m-1})$ is increasing in t, and under Assumption 4 $F(t|x, s)$ is FOSD-decreasing in x.

That is, for any s and $x'' > x'$, $r(x'', s) < r(x', s)$. Therefore $p^*(x', s) \geq p^*(x'', s)$, for any s (with strict inequality for some capacities). If $\mathscr{P}^*$ involves some randomization due to tie-breaking, this inequality is in terms of FOSD of the distributions $\mathbb{P}^*(p|x, s)$, and the support of $\mathbb{P}^*(p|x, s)$, as a function of (x, s), can be compared in the strong set order.

Hence for any position p,

$$\bar{s}_p(x'') \equiv \min\{s_i : \mathbb{P}^*(p|x'', s_i) > 0\} \geq \bar{s}_p(x') \equiv \min\{s_i : \mathbb{P}^*(p|x', s_i) > 0\},$$

hence $\bar{s}_p(x)$ is increasing in x, for any p. □

Proposition 3

Proof. Suppose $v(t,p)$ has convex differences in t (the proof with concave differences is analogous). By Theorem 3, under Assumption 5, $r(x'',s) \geq r(x',s)$ for all s. Therefore we have to consider 2 cases.

First, if $\mathscr{P}^*$ does not involve randomization due to tie-breaking, because $\mathscr{P}^*$ is monotonic in $r(x,s)$, we have that $p^*(x'',s) \geq p^*(x',s)$, with a strict inequality for some capacities and realizations of the observables.

Second, if $\mathscr{P}^*$ involves randomization due to tie-breaking, since (x'',s) and (x',s) are not payoff-equivalent, it must be the case that $|\operatorname{supp}[p^*(x'',s)] \cap \operatorname{supp}[p^*(x',s)]| \leq 1$, and moreover $\min\{\operatorname{supp}[p^*(x'',s)]\} \geq \max\{\operatorname{supp}[p^*(x',s)]\}$, and $\operatorname{supp}[p^*(x'',s)] \geq \operatorname{supp}[p^*(x',s)]$ in the strong set order. Then $p^*(x'',s)$ FOSD-dominates $p^*(x',s)$. Thus in both cases we have that $p^*(x'',s) \geq p^*(x',s)$, which proves the first part of the proposition.

Finally, consider the sets of agents assigned to any position p from each category. Since $p^*(x'',s) \geq p^*(x',s)$ for each s, we have that $\{s : \mathbb{P}^*(p|x'',s) > 0\} \leq \{s : \mathbb{P}^*(p|x',s) > 0\}$ in the strong set order, for any p. □

Proposition 4

Proof. For sufficiency, assume categories are strongly comparable and suppose, for the sake of a contradiction, that in equilibrium the assignment of agents to positions is not monotonic with respect to the benefit index $r(x,s)$. Then there exists some agent l with benefit index $r(x_l,s_l)$ who is assigned to a position k and some agent h with higher benefit index $r(x_h,s_h) > r(x_l,s_l)$ who is assigned to a lower position $j < k$.

Note that if $x_l = x_k$, this implies a contradiction, because in this case $s_l < s_h$, since $r(x,s)$ is increasing in s, following Proposition 1. Therefore, since $\mathbb{E}[v(t,p)|x,s]$ is increasing in s, we have $\mathbb{E}[v(t,p)|x_l,s_l] < \mathbb{E}[v(t,p)|x_h,s_h]$, and thus decision-maker k can accept agent h instead of agent l and strictly increase $\sum_{i \in A_k} \mathbb{E}[v(t_i,p_k|x_i,s_i)]$, implying that the original assignment cannot be an equilibrium outcome, a contradiction.

On the other hand, if $x_l \neq x_h$, then strict comparability implies that either (i) there exists some signal s'_h such that $\mathbb{E}[v(t,p)|x_h,s'_h] = \mathbb{E}[v(t,p)|x_l,s_l]$ and $r(x_h,s'_h) = r(x_l,s_l)$, in which case $\mathbb{E}[v(t,p)|x_l,s_l] = \mathbb{E}[v(t,p)|x_h,s'_h] < \mathbb{E}[v(t,p)|x_h,s_h]$, analogously to the case above, because $r(x,s)$ and $\mathbb{E}[v(t,p)|x,s]$ are increasing in s, and therefore again decision-maker k can accept agent h instead of agent l and strictly increase $\sum_{i\in A_j}\mathbb{E}[v(t_i,p_j|x_i,s_i)]$, a contradiction; or (ii) $\forall\, s'_h$ we have $\mathbb{E}[v(t,p)|x_l,s_l] \le \mathbb{E}[v(t,p)|x_h,s'_h]$, in which case again, analogously, $\mathbb{E}[v(t,p)|x_l,s_l] \le \mathbb{E}[v(t,p)|x_h,s'_h] < \mathbb{E}[v(t,p)|x_h,s_h]$ and thus decision-maker k can accept agent h instead of agent l and strictly increase $\sum_{i\in A_j}\mathbb{E}[v(t_i,p_j|x_i,s_i)]$, a contradiction.

For necessity, it is enough to provide a counter-example where categories are not strongly comparable and some agent h with index $r(x_h,s_h)$ is in equilibrium assigned to a position j, while another agent l with lower index $r(x_,s_l) < r(x_h,s_h)$ is assigned to a higher position $k > j$. Consider the example with 2 positions, $\{p_1,p_2\}$, each with a capacity of $k_j = 1$, 2 distinct categories, x_1 and x_2, and 2 agents with signals s_1 and s_2, with values given by:

	p_1	p_2
$\mathbb{E}[v(t,p)\|x_1,s_1]$	4	5
$\mathbb{E}[v(t,p)\|x_2,s_2]$	1	3

Clearly x_1 and x_2 are comparable (since the positions are binary), but are not strictly comparable, and moreover $r(x_2,s_2) > r(x_1,s_1)$, and yet decision-maker 2 would prefer to accept agent 1 in any equilibrium, since $\mathbb{E}[v(t,p_2)|x_1,s_1] > \mathbb{E}[v(t,p_2)|x_2,s_2]$, and agent 1 would accept p_2 over p_1. □

Proposition 5

Proof. From Proposition 1 it follows that $\mathbb{E}[v(t,p)|x,s]$ is increasing in s for any p, and moreover $\mathbb{E}[v(t,p_2)|x,s] - \mathbb{E}[v(t,p_1)|x,s]$ is also increasing in s. Consider any category x and signal s. For any category x', since $\mathbb{E}[v(t,p_2)|x',s] - \mathbb{E}[v(t,p_1)|x',s]$ is continuous and increasing in s, either $\exists s'$ such that $\mathbb{E}[v(t,p_2)|x',s'] - \mathbb{E}[v(t,p_1)|x',s'] = \mathbb{E}[v(t,p_2)|x,s] - \mathbb{E}[v(t,p_1)|x,s]$, or $\mathbb{E}[v(t,p_2)|x',s'] - \mathbb{E}[v(t,p_1)|x',s'] > \mathbb{E}[v(t,p_2)|x,s] - \mathbb{E}[v(t,p_1)|x,s]$ for all s', or $\mathbb{E}[v(t,p_2)|x',s'] - \mathbb{E}[v(t,p_1)|x',s'] < \mathbb{E}[v(t,p_2)|x,s] - \mathbb{E}[v(t,p_1)|x,s]$ for all s'. In all 3 cases, the condition for comparability is satisfied for all p,p', hence x and x' are comparable. □

Proposition 6

Proof. Let $v(t,p) = v_1(t) \cdot v_2(p)$ and consider any family of signal structures $\{F_{t,x}\}$. Consider two categories, x and x', and an arbitrary signal s. Then

$$\mathbb{E}[v(t,p)|x,s] = \mathbb{E}[v_1(t)|x,s] \cdot v_2(p).$$

Consider $\{\mathbb{E}[v_1(t)|x',s] : s \in S\}$, i.e. the range of possible values $\mathbb{E}[v_1(t)|x',s]$ over all s. Because $\mathbb{E}[v(t,p)|x,s]$ is continuous in s, this set is an interval. There are two cases.

Case 1: $\mathbb{E}[v_1(t)|x,s] \in \{\mathbb{E}[v_1(t)|x',s] : s \in S\}$. Then $\exists s' \in S$ s.t. $\mathbb{E}[v_1(t)|x,s] = \mathbb{E}[v_1(t)|x',s']$. Hence $\exists s'$ s.t. $\forall p$:

$$\mathbb{E}[v(t,p)|x,s] = \mathbb{E}[v_1(t)|x,s] \cdot v_2(p) = \mathbb{E}[v_1(t)|x',s'] \cdot v_2(p) = \mathbb{E}[v(t,p)|x',s'].$$

Case 2: $\mathbb{E}[v_1(t)|x,s] \notin \{\mathbb{E}[v_1(t)|x',s] : s \in S\}$. If $\mathbb{E}[v_1(t)|x,s] <$ [resp. $>$] $\{\mathbb{E}[v_1(t)|x',s] : s \in S\}$, then $\forall s'$ and $\forall p$ we have

$$\mathbb{E}[v(t,p)|x,s] = \mathbb{E}[v_1(t)|x,s] \cdot v_2(p) < \text{[resp. >]}\, \mathbb{E}[v_1(t)|x',s'] \cdot v_2(p) = \mathbb{E}[v(t,p)|x',s']$$

In both cases we conclude that x and x' are strongly comparable. □

Proposition 7

Proof. Consider two categories, x and x', and an arbitrary signal s.

$\{F_{t,x}\}$ and $\{F_{t,x'}\}$ are equivalent, so $\exists \delta$ s.t.

$$\widetilde{w}_{p,x}(v) = \widetilde{w}_{p,x'}(v) + \delta$$

for any p, for all $v \in \text{Im}[\widetilde{v}_{p,x}(s)] \cap \text{Im}[\widetilde{v}_{p,x'}(s)]$.

There are 2 cases to consider.

Case 1: $\text{Im}[\widetilde{v}_{p,x}(s)] \cap \text{Im}[\widetilde{v}_{p,x'}(s)] = \varnothing$. Because $\widetilde{v}_{p,x}(s)$ and $\widetilde{v}_{p,x'}(s)$ are continuous, the intermediate value theorem implies that either $\widetilde{v}_{p,x}(s) > \widetilde{v}_{p,x'}(s')$ $\forall s, s'$, or $\widetilde{v}_{p,x}(s) < \widetilde{v}_{p,x'}(s')$ $\forall s'$. This holds for any arbitrary p, hence part 2 of the definition of strong comparability is satisfied.

Case 2: $\text{Im}[\widetilde{v}_{p,x}(s)] \cap \text{Im}[\widetilde{v}_{p,x'}(s)] \neq \varnothing$. Take any $v \in \text{Im}[\widetilde{v}_{p,x}(s)] \cap \text{Im}[\widetilde{v}_{p,x'}(s)]$, and take δ s.t.

$$\widetilde{w}_{p,x}(v) = \widetilde{w}_{p,x'}(v) + \delta.$$

Let $s' = s - \delta$. Then

$$\mathbb{E}[v(t,p)|x,s] = \mathbb{E}[v(t,p)|x',s-\delta] = \mathbb{E}[v(t,p)|x',s']$$

holds $\forall v \in \text{Im}[\widetilde{v}_{p,x}(s)] \cap \text{Im}[\widetilde{v}_{p,x'}(s)]$. Hence part 1 of the definition of strong comparability is satisfied for all such v.

Next, consider $v < \text{Im}[\widetilde{v}_{p,x}(s)] \cap \text{Im}[\widetilde{v}_{p,x'}(s)]$. Because $\widetilde{v}_{p,x}(s)$ is continuous and increasing in s, this implies that $\forall s' \in S$ we have $\mathbb{E}[v(t,p)|x',s'] > \mathbb{E}[v(t,p)|x,s]$ or $\mathbb{E}[v(t,p)|x',s'] < \mathbb{E}[v(t,p)|x,s]$, $\forall p$. Analogously, the same holds for $v > \text{Im}[\widetilde{v}_{p,x}(s)] \cap \text{Im}[\widetilde{v}_{p,x'}(s)]$. Hence part 2 of the definition of strong comparability is satisfied for all such v.

In both cases we conclude that x and x' are strongly comparable. $\square$