Intersectionality and Individual Decision Making

Caroline W. Liqui Lung *

Version: June 14, 2022

Intersectionality refers to the interconnection of social organizations, such as gender and race, that create interdependent systems of disadvantage. This paper introduces intersectionality in individual decision making, analyzing the phenomenon in a setting without interaction between agents. I extend the model presented in Liqui Lung (2022) to analyze agents with multidimensional social identities. I show how an intersectional lens sheds light on choice behavior and inequalities that are not visible when using a one-dimensional perspective on social identity. I discuss the effects of social constraints, such as stigmatization, and use the framework to analyze the effects of affirmative action policy on choice behavior. I show how an intersectional view leads to novel insights that are important for the development of the adequate policies to achieve diversity and fight harmful stereotypes.
(JEL: D-81, D-91, Z-13)

*Paris School of Economics (PSE, ENPC) (Email: caroline.liqui-lung@psemail.eu). I would like to thank Olivier Compte, Philippe Jehiel, Roland Benabou, Alessandra Casella and Willemien Kets for their useful comments and suggestions.

1 Introduction

The term ‘intersectionality’ was first introduced by Kimberlé Crenshaw to discuss the fact that discrimination on the labor market towards women and people of color interacts, disadvantaging women of color more than white women or men of color (Crenshaw, 1991). The concept is now more broadly used, and the Oxford dictionary defines the term as ‘the interconnected nature of social organizations, such as race, class, age and gender creating overlapping and interdependent systems of discrimination and disadvantage’.

The inequalities associated to intersectionality that are currently studied are always induced by some type of interaction between agents, resulting in discrimination and differences in material benefits and socio-economic opportunities (Yuval-Davis (2006), Cassan and Vandewalle (2021), Hughes (2011)). This paper shows how intersectionality also plays a role in individual decision making, a setting in which there is no interaction between agents. Specifically, I extend the model in Liqui Lung (2022) to analyze how social context affects belief formation and choice behavior of agents with multidimensional social identities. I show how an intersectional view on social groups sheds light on novel insights, and discuss how these insights are important for the development of the policies necessary to fight harmful stereotypes and achieve diversity.

The paper analyzes a model in which the effect of social context on belief formation is determined endogenously, without assuming that social context directly affects utility or preferences, nor that social context contains any relevant information for decision making. Agents choose between undertaking a *Competitive* task or a *Non-Competitive* outside option. The probability of success of the outside option is known and the same for all agents. The *Competitive* task has a probability of success that depends on individual-specific characteristics. The key assumption in the model is that agents only have a noisy perception of their individual-specific probability of success.

The novelty in this paper is that agents are described by a social type that is composed of multiple observable characteristics, that can represent for example an agent’s gender, race or social class. Agents observe social identity cues that stem from the prevalence of a certain subgroup among the successful individuals in social context. These subgroups can be based on a single observable characteristic or a social type. Because all characteristics are independently distributed from the individual-specific probability of success of the *Competitive* task, these social identity cues are irrelevant in a Bayesian sense. I nevertheless envision agents that have an imperfect idea about their environment, and think that using their social context could be useful to form a belief about their individual-specific probability of success. Agents choose from a family of subjective belief-formation processes. These processes are either based solely on their noisy perception, or are such that their noisy perception is influenced by a certain social identity cue in a direction contingent on the agent's observable type. In other words, when an agent belongs to the socially more successful group according to a certain cue, this cue leads to an optimistic interpretation of her noisy perception in belief formation, while vice versa for agents belonging to the socially less successful group.

Agents derive utility from being successful, and they choose a task to maximize subjective expected utility. The choices of subjective belief formation induce choices of tasks, and, at the aggregate level, these choices of tasks give rise to social identity cues. I use a static solution concept to study the mutually stable choices of subjective beliefs and tasks in this dynamic process. Specifically, agents choose their belief formation according to an individual optimality criterion that states that a subjective belief formation process is optimal when it maximizes expected utility on average over all possible realizations of the agent's noisy perception. Consequently, I analyze the existence and stability of the fixed points in the social identity cues induced by these individually optimal strategies.

The main idea of the model is that the multidimensionality of social identity increases the set of possible belief formation processes. Furthermore, agents choose their belief-formation process according to a fitness criterion. This implies that agents have a certain degree of control to focus on specific dimensions of their social identity, while possibly ignoring another dimension[1]. I analyze the behavior of agents in two different settings. In the first setting, agents are not constrained in their choice of social identity cues in belief formation. In the second setting, I analyze the effects of various social constraints that limit an agent's ability to ignore certain dimensions of her social identity.

As in Liqui Lung (2022), social identity cues that are irrelevant in a Bayesian sense, become useful when their use in belief formation generates a bias towards the welfare-maximizing choice of task. Additionally, I show that, when multiple social identity cues have this property, agents should use the cue that generates the strongest bias in the correct direction. This sheds light on how agents endogenously determine which dimensions of social identity affect belief formation as a function of their exogenously specified social type, their social context and their underlying ability.

The idea that agents choose to focus on a certain dimension of social identity to increase utility also appears in other decision-making settings. Atkin et al. (2021) shows how ethnic and religious identities in India are determined by group status, group salience

[1]This idea is related to the concept of 'agency' in sociology. Agency is defined as a situated form of subjective action characterized by the encounter, on the one hand, between individual intentionality, value orientations, aspirations and resistances, capacity to give meaning to action and definitions of reality, and on the other hand, the constraints of the social and material environment (Rebughini, 2021).

and the market cost of following a group's prescribed behaviors. Shayo (2009) discusses how group identification is driven by group status and perceived similarity along the different dimensions of social identities, where people prefer groups with a higher social status. Jia and Persson (2019) shows how the choice of ethnicity for children in ethnically mixed marriages in China is driven by the interaction between material benefits that can only be received when belonging to certain minority groups, and existing social norms of following the father's identity. People also manipulate their identity when this works to their advantage. Qian and Nix (2015) shows how the rate at which black Americans were 'passing' as white was correlated with geographical relocation to communities with higher percentages of whites, and with better political, economic and social opportunities for whites relative to blacks. Cassan (2015) similarly shows how the Punjab alienation of land act led to a movement of identity-manipulation

Because of the noisy perception of their individual-specific probability of success, agents can make two types of mistake. First, they may choose to undertake the *Competitive* task, while this is not their welfare-maximizing choice. Second, they may choose to not undertake the *Competitive* task, while this is in fact their welfare-maximizing choice. The key insight of this paper is that we can distinguish two categories of social types. First, there are *mixed* social types, who belong to the socially more successful group according to one observable characteristic, while they belong to the socially less successful group according to the other characteristic. Secondly, there are *one-sided* social types, who belong to either the socially more or less successful group along the lines of both observable characteristics. Agents with the latter social type belonging to the socially less successful group can only bias their noisy perception downwards, while agents with a *one-sided* social type belonging to the socially more successful group can only bias their noisy perception upwards. Agents with *mixed* social types have the option to bias their noisy perception both upwards and downwards. They can therefore decrease the likelihood of making both types of mistake. Consequently, agents with *mixed* social types have on average a higher expected utility than agents with *one-sided* social types.

I use two approaches to analyze choice behavior at the aggregate level. A *One-dimensional lens* aggregates agents along the lines of one observable characteristic. An *Intersectional lens* aggregates agents along the lines of social types. A *One-dimensional* lens shows that agents belonging to the socially more or less successful group according to a single dimension of social identity have on average a similar potential to improve decision making. They are nevertheless more likely to make a different type of mistake. This asymmetry drives differences in choice behavior between these two social groups. An *Intersectional lens* shows nevertheless that not all agents within a *one-dimensional* subgroup have the same potential to improve decision making. Specifically, while agents with a *mixed* social type have a similar probability to choose the *Competitive* task, this probability is larger for agents with the socially more successful *one-sided* social type, while smaller for agents with the socially less successful *one-sided* social type. An *Intersectional lens* sheds therefore light on inequalities that are not visible with a *One-dimensional lens*, and shows that differences in choice behavior along the lines of socially more and less successful subgroups are driven by agents with a *one-sided* social type.

Consequently, I show the existence of a population equilibrium in which a priori identical individuals with different social types end up choosing different tasks. Differences in choice behavior across subgroups can therefore be persistent. When this is the case, agents with a *one-sided* social type that belong to the socially more successful group are most overrepresented among the successful individuals, while agents with a *one-sided* social type that belong to the socially less successful group are most underrepresented.

The relative advantage that agents with *mixed* social types have over agents with *one-sided* social types can disappear when agents are constrained in their choice of social identity cue. The first type of constraint I analyze is such that agents are unable to ignore one dimension of their social identity in belief formation, which I refer to as 'stigmatization'. Agents with a *mixed* social type can now no longer correct for both types of mistakes. This brings their potential to improve decision making to a similar level as the potential of agents with *one-sided* social types. Furthermore, this constraint increases differences in choice behavior along the lines of the stigmatized dimension of social identity, while it decreases the differences in choice behavior along the lines of the non-stigmatized dimension. The second type of constraint I consider is such that agents cannot ignore any dimension of their social identity in belief formation. This constraint captures a setting in which agents are only able to identify with agents having the same social type[2]. Similarly, this constraint takes away the ability of agents with a *mixed* social type to potentially correct for both two types of mistake.

Finally, I use the framework to analyze the effects of affirmative action policy on choice behavior. I show how a quota targeting one dimension of social identity can have the undesirable effect of increasing differences in choice behavior along the lines of other dimensions of social identity. Moreover, in the face of social constraints, a quota may have very little effects. This is in line with the empirical results of for example Hughes (2011), Folke et al. (2015) and Tan (2014). On the other hand, when the quota affects agents with *one-sided* social types, this can reduce differences in choice behavior along

[2]This point is raised by Crenshaw (1991), who writes about how black women often have difficulty to identity with women in general, specifically white women.

multiple dimensions of social identity simultaneously. This is in line with Cassan and Vandewalle (2021), who show how a gender quota in India affects the representation along the gender and caste dimension simultaneously, because women's political participation is correlated with belonging to a low caste. Finally, I show how affirmative action policy could hurt individual decision making and decrease aggregate welfare, and suggest how this policy could be complemented with informational policies.

The structure of the paper is as follows. Section 2 presents the model[3]. Section 3 presents the results in the absence of any constraints, while section 4 discusses the effects of different types of social constraints. Section 5 analyzes the effects of affirmative action policy. Finally, section 6 concludes. All formal proofs can be found in the Appendix.

2 The Model

2.1 The Environment

I consider a society with $i = 1, ..., N$ agents, with N arbitrarily large. Each agent i chooses an action $a_i \in \{C, NC\}$, where C and NC represent classes of tasks of a *Competitive* and a *Non-Competitive* type. The outcome of a_i can be either *'success'* or *'failure'* and is represented by the variable $Y_i \in \{1, 0\}$. The probability of success for a *Competitive* task depends on an agent's individual characteristics. This probability is represented by the continuous variable $\alpha \in [0, 1]$, and is distributed over the population following a distribution f_α. For each agent i, the probability of a successful outcome $Y_i = 1$ conditional on choosing the *Competitive* task is fixed and given by,

$$p(Y_i = 1 | a_i = C) = \alpha_i \tag{1}$$

The *Non-Competitive* task has a probability of success $\gamma \in [0, 1]$ that is known and the same for all agents. Therefore, for all i,

$$p(Y_i = 1 | a_i = NC) = \gamma \tag{2}$$

More generally, γ can be interpreted as the attractiveness of the *Non-Competitive* task relative to the *Competitive* task. Agents make a series of choices during their lifetime between *Non-Competitive* tasks and *Competitive* tasks of which the probability of success depends on similar individual characteristics. For example, α_i may be related to a person's mathematical ability or leadership qualities. Although it is realistic to assume

[3]The exposition of the model follows Liqui Lung (2022), where I extend this model by introducing multi-dimensional social identities.

that α and γ vary slightly from choice set to choice set, the main insights of this model can also be transmitted with a simpler and more parsimonious model. Therefore, in the following, I assume that the variables α_i and γ are fixed over an agent's lifetime[4].

Noisy Belief Formation - The main assumption in the model is that α_i is unobservable, and agents only have a noisy perception or belief $\hat{\alpha}_i$ regarding their own probability of success α_i. I assume this belief stems from a noisy inference process that is unbiased, but always incomplete. Therefore, I pose that $\hat{\alpha}_i$ stems from a distribution g_{α_i} with $E(\hat{\alpha}_i) = \alpha_i$. The assumption that the noisy belief is unbiased can be challenged, but the objective of this assumption is to show that a systematic bias in the individual-specific belief formation process is not the mechanism that drives the results in this model.

The Social Context - Each agent is described by multiple *observable characteristics* that can represent a wide array of items, such as the agent's gender, ethnicity, social class or age. These observable characteristics are public information, meaning that both the agent herself and all other agents in the society can observe them. To simplify to exposition of this model, I assume each agent has a social type $\Theta_i = (\theta_i^k)_{k \in \{A,B\}}$, which is a vector of two binary[5] observable characteristics $\theta_i^k \in X = \{0, 1\}$. Let $x \in \{0, 1\}$ denote a possible realization of θ_i^k, and let $t \in T = \{11, 10, 01, 00\}$ be a possible realization of the social type Θ_i. Consequently, p_x^k is the fraction of the population with observable characteristic $\theta^k = x$, and p_t is the fraction of the population with social type $\Theta = t$. Each agent i is fully described by (α_i, Θ_i). To isolate the mechanism through which social identity affects choice behavior in this model, I assume that the probabilities α_i and the social types Θ_i are independently distributed over the population. Furthermore, to simplify the exposition of the model, I also assume that the observable characteristics θ_i^k are independently distributed over the population.[6]

Agents have access to public data. This public data consists of the outcome variables and the observable characteristics of other agents that have made a similar choice at some earlier point in time. Society typically structures public information. For the exposition of the model, I assume agents focus on the fraction of successful individuals in the *Competitive* task that have an observable characteristic, or social type in common, among all successful individuals in this task[7].

[4]See Liqui Lung (2022) Section 5.2 for a discussion regarding the consequences of this simplification.
[5]The model can also account for non-binary observable characteristics.
[6]See Section 3.4 for a discussion on the implications of correlated observable characteristics.
[7]See Liqui Lung (2022) for a discussion about the effects of different structures on information.

Let $\mathcal{N}^k_{C,x} = \{i \in N, \theta^k_i = x, a_i = C\}$ be the set of agents with $\theta^k_i = x$ that have chosen the *Competitive* task. Similarly, let $\mathcal{N}_{C,t} = \{i \in N, \Theta_i = t, a_i = C\}$ be the set of agents with social type $\Theta_i = t$ that have chosen the *Competitive* task. Finally, let $\mathcal{N}_C = \{i \in N, a_i = C\}$ be the set of all agents that have chosen the *Competitive* task, which implies $\mathcal{N}^k_{C,x}, \mathcal{N}_{C,t} \subset \mathcal{N}_C$. Society then provides the statistics,

$$\pi^k_x = \frac{\sum_{i \in \mathcal{N}^k_{C,x}} Y_i}{\sum_{i \in \mathcal{N}_C} Y_i} \qquad \qquad \pi_t = \frac{\sum_{i \in \mathcal{N}_{C,t}} Y_i}{\sum_{i \in \mathcal{N}_C} Y_i}$$

for all $k \in \{A, B\}$, $x \in \{0, 1\}$ and $t \in \{11, 10, 01, 00\}$. These are the fractions of successful individuals with characteristic $\theta^k_i = x$ or social type $\Theta_i = t$ among all successful individuals that have chosen the *Competitive* task[8]. I call the fractions π^k_x and π_t the '*social identity cues*' for agents with social type $\Theta_i = (\theta^A_i, \theta^B_i)$. Furthermore, I will refer to the fractions π^k_x as '*one-dimensional social identity cues*', while I will refer to the fractions π_t as '*two-dimensional social identity cues*'. The social context of the population is defined as the vector $\Pi = (\pi_t)_{t \in T}$. Because α_i, Θ_i and θ^k_i are independently distributed over the population, this social context contains no relevant information about the individual-specific probability of success when undertaking a *Competitive* task. Instead, I introduce the option to agents to bias their noisy belief $\hat{\alpha}_i$ using this public data.

Subjective Belief Formation - I model agents that have an imperfect idea about their economic environment and think that using their social context could possibly be useful to form a belief about their probability of success α_i. Consequently, I introduce a family of belief formation processes with which agents form a subjective belief $\hat{p}_i$ about their probability of success of a *Competitive* task α_i, and I assume agents have some discretion in finding out which belief formation process suits them best. Specifically, I have the following story in mind. Assume that agents have a natural 'urge' to look at others like them when they are not sure what to do, and people have the option to either *Repress* or *Not Repress* this urge. Furthermore, agents can decide whether they consider others like them to be defined as agents with their same observable characteristic θ^A_i, their observable characteristic θ^B_i or as agents with their entire social type Θ_i.

[8]The fact that agents only consider the successful individuals that have chosen the task captures the survivors bias (Denrell, 2003). This bias does not drive the results.

To simplify the exposition of the model, I first focus on agents that can only define other like them along the lines of a single characteristic[9]. Consequently, agents choose a strategy $\sigma_i \in \{A, B, R\}$ that results in a belief about α_i equal to $\hat{p}_i^{\sigma_i} \in \{\hat{p}_i^A, \hat{p}_i^B, \hat{p}_i^R\}$. Specifically, let η be a '*response function*' that is non-decreasing, such that,

$$\eta(\pi, p) = \begin{cases} > 1 & \text{if } \pi > p \\ 1 & \text{if } \pi = p \\ < 1 & \text{if } \pi < p \end{cases} \tag{3}$$

and let $\eta_{k,x} \equiv \eta(\pi_x^k, p_x^k)$. Then,

$$\hat{p}_i^{\sigma_i} = \begin{cases} \hat{\alpha}_i & \text{if } \sigma_i = R \\ \eta_{k,x}\hat{\alpha}_i & \text{if } \sigma_i \in \{A, B\} \end{cases} \tag{4}$$

In other words, depending on whether agents let their belief-formation process be influenced by social context, their subjective belief can take three values[10]. With a subjective Bayesian interpretation in mind, $\sigma_i = R$ corresponds to a world view in which private and observable characteristics are uncorrelated, while $\sigma_i \neq R$, corresponds to a world view in which private and observable characteristics are correlated, with (π_x^k, p_x^k) informing about the sign and strength of that correlation[11]. When choosing $\sigma_i \neq R$, the agent biases her noisy belief $\hat{\alpha}_i$ in the direction contingent on her social type. When the agent's chosen subgroup is overrepresented among the successful individuals at the *Competitive* task in the society, the use of social identity cues in the belief-formation process leads to an optimistic interpretation of the noisy belief $\hat{\alpha}$, while this leads to a pessimistic interpretation of the noisy belief when the agent's chosen subgroup is underrepresented among the successful individuals in the society[12].

Subjective Utility Maximization - Agents derive utility from being successful. The utility function can therefore be represented by $u_i = Y_i$. Each agent chooses her action a_i to maximize $E(u_i)$ given her subjective belief $\hat{p}_i^\sigma$. Consequently, she chooses the *Competitive* task if and only if $\hat{p}_i^\sigma > \gamma$. One could say therefore that agents are subjectively

[9]Section 3.3 discusses the effects of agents also using *two-dimensional* social identity cues.

[10]I don't model the agent's exact thought process leading to the possible beliefs. The objective is not to propose a particular functional form, nor to root it in a specific subjective Bayesian model, but to investigate how properties of the response function are conductive to the phenomenon I mean to describe.

[11]In this case, the model could be interpreted as agents making an attribution error.

[12]In the specific case where the social identity cues are equal to the corresponding population fractions, all strategies are equivalent.

rational given the process that determines their subjective beliefs. Furthermore, the model allows for two different interpretations. One interpretation is that the instrument σ_i mechanically alters the agents' subjective belief $\hat{p}_i^{\sigma}$, where $\hat{p}_i^{\sigma_i} \in \{\hat{p}_i^A, \hat{p}_i^B, \hat{p}_i^R\}$. Another interpretation is that agents have the option to use the social identity cue to alter choice in a direction contingent on their observable type. Formally, subjective expected utility maximization implies that the agent is effectively comparing thresholds, such that agent i chooses $a = C$ if and only if $\hat{\alpha}_i > \gamma_i$, where

$$\gamma_i = \begin{cases} \gamma & \text{when } \sigma_i = R \\ \frac{\gamma}{\eta_{k,x}} & \text{when } \sigma_i \in \{A, B\} \end{cases} \tag{5}$$

The use of the social identity cue in the belief-formation process implies that the agent inflates or deflates the threshold for $\hat{\alpha}$ above which she thinks she is 'good enough' to undertake the *Competitive* task. The strategy set can therefore also be directly specified as the choice set in terms of γ_i. This choice set is different for agents with different social types, which will be the key driver of the equilibrium results.

2.2 The Solution Concept

Whether agents will let social context affect their beliefs will affect their choice of task. This choice behavior will lead to outcomes that induce social identity cues. These social identity cues again affect the way agents form subjective beliefs. To tractably capture the fixed points of such a dynamic process, I use a static solution concept in which I assume that, given a social context, agents choose their strategy σ according to its fitness value. This fitness is determined by an individual optimality criterion I will specify and justify below. I then define a population equilibrium as a fixed point in the social context that is induced by individually optimal strategy choices. This solution concept is in line with the view that the optimal choice of the strategy σ arises from a learning process that operates faster than the dynamics in the social context, where the learning of the optimal strategy happens during the lifetime of an agent through her experience with similar tasks, while changes in the social context of a specific task arise from agents belonging to different generations making this specific choice of task once in their lifetime.

Individual Optimality - Let $\Phi_{\alpha,t,\sigma_i,\Pi} = P(a = C|\alpha, t, \sigma_i, \Pi)$ be the induced probability that an agent with α and social type $\Theta_i = t$ playing strategy σ_i given a social context Π chooses the *Competitive* task. Then,

$$\Phi_{\alpha,t,\sigma_i,\Pi} = P(\hat{p}_i^\sigma > \gamma|\alpha) \tag{6}$$

This probability Φ is determined objectively conditional on the choice of strategy σ_i. From an outsiders perspective, the expected pay-off for agent i with α_i and $\Theta_i = t$ playing σ_i given Π over all possible realizations of $\hat{\alpha}$ is,

$$V_i(\sigma_i) = \alpha\Phi_{\alpha,t,\sigma_i,\Pi} + \gamma(1 - \Phi_{\alpha,t,\sigma_i,\Pi}) \tag{7}$$

with $\sigma_i \in \{A, B, R\}$. Individual optimality can then be defined as follows.

DEFINITION 1 (Individual Optimality): *The strategy σ_i^* is optimal for the agent from an individual perspective when,*

$$\sigma_i^* = \operatorname*{argmax}_{\sigma_i} V_i(\sigma_i)$$

Individual optimality means that an agent uses her social identity cue to maximize her expected pay-off on average over all possible realizations of $\hat{\alpha}_i$. The fitness value of a strategy σ is therefore determined by an agent's type (α, Θ) and the social context Π. I assume that agents can compare $V_i(R), V_i(A)$ and $V_i(B)$ and choose their strategy σ_i according to the individual optimality criterion. This assumption can be justified with the view that agents have learned their optimal strategy from their own experience with similar tasks in a similar social context through for example reinforcement learning or a sampling process. The true probability α_i determines the outcomes the agent observes in this process, which enables her to learn whether it is optimal to use social identity cues without precise knowledge of the relationship between her choice of strategy, choice of task and the observed outcome[13]. One could say that agents are boundedly rational in the sense that not all belief formation processes can be compared, meaning not all possible functions of $\hat{\alpha}$ and π_x^k. This aspect of bounded rationality should be considered as a modelling device that helps to keep the model parsimonious[14].

[13]The idea is that, although I make the simplifying assumption that α and γ are fixed throughout the lifetime of an agent, for each choice, agents receive a new realization of the belief $\hat{\alpha}_i$.

[14]Because of the simplifying assumption that α and γ are fixed, if agents could compare all such belief formation processes, they would behave as a Bayesian and choose $a = C$ when $\alpha > \gamma$. Because α, Θ and the observable characteristics are independently distributed, a Bayesian analysis would be degenerate in

Population Equilibrium - Let σ be the collection of σ_i. Because N is arbitrarily large, each collection of strategies σ and social context Π generate choices and successes that in turn generate public data $\tilde{\Pi}$ such that,

$$\tilde{\pi}_t(\sigma, \Pi) = \frac{p_t \int \alpha \Phi_{\alpha,t,\sigma,\Pi} f(\alpha) d\alpha}{\sum_{t \in T} p_t \int \alpha \Phi_{\alpha,t,\sigma,\Pi} f(\alpha) d\alpha} \tag{8}$$

where $f(\alpha)$ is the probability density function of α and $\tilde{\pi}_t(\sigma, \Pi)$ is the social identity cue induced by strategies σ and a social context Π. An equilibrium in the model can now be defined as follows.

DEFINITION 2 (Population Equilibrium): *A pair of strategies and a social context* $\{\sigma, \Pi\}$ *constitutes a population equilibrium, when* $\sigma = \sigma^*$ *for all agents given* Π*, and when* Π *is such that,*

$$\Pi = \widetilde{\Pi}(\sigma, \Pi) \tag{9}$$

A population equilibrium arises therefore when all agents play their optimal strategy given their social context, and when these strategy choices induce a fixed point in the social context.

3 Results

3.1 Multidimensional Identities and Individual Choice Behavior

Agents choose whether social context influences belief formation. Furthermore, when they let their social identity cues play a role, they choose to focus on the subgroup that either has their observable characteristic θ^A, or their characteristic θ^B in common. In the following example, I illustrate what belief formation looks like when agents choose their strategy according to the individual optimality criterion specified in Definition 1.

Example - Consider a high school in which students choose to enter a math competition (C) or a competition that tests a mix of abilities (NC). Students observe the previous pool of successful students in this competition, and see their gender and whether they are Asian or not. Let gender be represented by $\theta^A \in \{1, 0\}$, where $\theta^A = 1$ represents being male. Let being Asian or not be represented by $\theta^B \in \{1, 0\}$, where $\theta^B = 1$ represents

this case. The model shows therefore the difference with a Bayesian model, by analysing whether, when agents are not able to draw all inferences given the structure of the model, this can open the door for agents to use information that is irrelevant, but that could still improve decision making.

being Asian. Furthermore, for simplicity, assume that $p_t = p_{t'}$ for all $t, t' \in T$. Finally, let the current social context be such that $\pi_0^k < \pi_1^k$ for $k \in \{A, B\}$ and $\pi_1^A = \pi_1^B$. In other words, Asian and male students are relatively overrepresented to the same degree in the previous cohort of successful students, while non-Asian and female students are relatively underrepresented to the same degree.

Figure 1 shows the choices of belief formation this social context offers to a non-Asian male student. On the axis, you can find the thresholds γ_i that follow from the strategies $\sigma_i \in \{A, B, R\}$, while the arrows show the probabilities $\Phi_{\alpha,t,\sigma_i,\Pi}$ with which the student enters the math competition induced by each strategy σ_i given the social context Π.

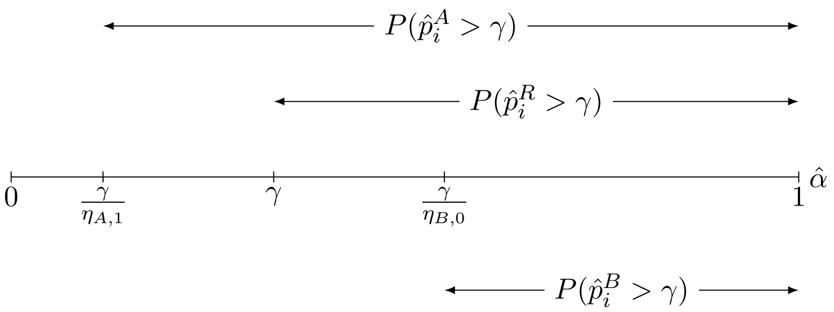

Figure 1: The probabilities to choose $a = C$ induced by $\sigma_i \in \{A, B, R\}$ for an agent with type $t = 10$

Consider first a non-Asian male student with $\alpha > \gamma$. To maximize expected utility, this student should enter the math competition. When he represses the use of social identity cues in belief formation, he will choose this option for all realizations of $\hat{\alpha} > \gamma$. When he uses social identity cues in belief formation, and focusses only on the non-Asian successful students from the previous cohort, he will choose to enter the competition for all realizations $\hat{\alpha} > \frac{\gamma}{\eta_{B,0}}$. Figure 1 shows that this decreases the likelihood he chooses his welfare-maximizing option. When he focusses on all successful male students, he will choose to enter the competition for all realizations $\hat{\alpha} > \frac{\gamma}{\eta_{A,1}}$. This strategy increases the probability $\Phi_{\alpha,t,\sigma_i,\Pi}$. According to the individual optimality criterion, a non-Asian male student with $\alpha > \gamma$ should therefore believe that being a male student increases his chances of success in the competition, while he should disregard the fact that non-Asian students are underrepresented among the successful students in the previous cohort.

On the other hand, a non-Asian male student with $\alpha < \gamma$ should not enter the math competition. The social identity cue based on all non-Asian students biases his own noisy perception in the welfare-maximizing direction. Therefore, according to the individual optimality criterion, a non-Asian male student with $\alpha < \gamma$ should believe that being non-Asian decreases his chances of success in a math competition, and disregard the fact that male students are overrepresented among those that were successful in the previous cohort. A similar reasoning can be applied to determine the optimal strategies of Asian male students, Asian female students and non-Asian female students.

Define $\overline{\eta}_t = \max_k \eta_{k,x}$ and $\underline{\eta}_t = \min_k \eta_{k,x}$. Similarly, let $\overline{\kappa}_t = \text{argmax}_k \, \eta_{k,x}$ and $\underline{\kappa}_t = \text{argmin}_k \, \eta_{k,x}$. Then, these results can be generalized as follows.

PROPOSITION 1 (Individually Optimal Belief Formation): *Consider an agent with social type Θ_i. For agents with $\alpha > \gamma$, $\sigma_i^* = \overline{\kappa}_t$ if and only if $\overline{\eta}_t > 1$. Otherwise, $\sigma^* = R$. For agents with $\alpha < \gamma$, $\sigma_i^* = \underline{\kappa}_t$ if and only if $\underline{\eta}_t < 1$. Otherwise, $\sigma^* = R$.*

Social identity cues give agents the option to nudge beliefs in a particular direction determined by social context. Proposition 1 shows that, when only using social identity cues when this biases beliefs towards the welfare-maximizing task, agents can improve decision making on average. Furthermore, when multiple social identity cues bias beliefs in the correct direction, it is optimal for agents to focus on the cue that induces the strongest bias in beliefs. This shows how, in a specific choice setting, agents endogenously determine which dimensions of social identity affect belief formation as a function of their exogenously specified social type, their social context and their underlying ability.

3.2 Aggregate Choice Behavior across Social Groups

The individual choice behavior induced by Proposition 1 can be aggregated in two different ways. First, we can aggregate this data while creating social groups along the lines of single observable characteristics, such as male versus female students. Secondly, we can aggregate this data while creating social groups along the lines of the different social types, resulting in Asian male students, Non-Asian male students, Asian female students and non-Asian female students. Definition 3 defines these different approaches as using a *One-dimensional lens* and an *Intersectional lens.*

DEFINITION 3: *The effect of social context on decision making at the aggregate level can be evaluated using the following approaches:*

- **One-dimensional lens**: *aggregating data by creating social groups defined along the lines of one observable characteristic* θ^k
- **Intersectional lens**: *aggregating data by creating social groups defined along the lines of social types* t

In the following, I evaluate the effects of the optimal individual choice behavior implied by Proposition 1 at the aggregate level, while aggregating data using these two different approaches. This enables me to show how using an *Intersectional Lens* sheds light on asymmetries and inequalities that are invisible when using a *One-Dimensional lens.*

3.2.1 Potential to Improve Decision Making

An agent's noisy perception $\hat{\alpha}_i$ of her true probability α_i can cause her to make two types of mistakes. Either, she chooses the *Non-Competitive* task while $\alpha > \gamma$, or she chooses the *Competitive* task while $\alpha < \gamma$. She could decrease the likelihood of making these mistakes by nudging her beliefs in the direction of her welfare-maximizing choice with the use of her social identity cues. The degree to which she is able to decrease this likelihood depends on the strength and direction of the bias induced by her possible strategies σ_i given the current social context Π. To capture this potential to improve decision making, I define the following measures for agents with $\Theta_i = t$. Let,

$$D_t^+ = \max(\overline{\eta}_t - 1, 0) \tag{10}$$

be the potential to increase the likelihood of choosing the *Competitive* task when $\alpha > \gamma$, and

$$D_t^- = \max(1 - \underline{\eta}_t, 0) \tag{11}$$

be the potential to increase the likelihood of choosing the *Non-Competitive* task when $\alpha < \gamma$. At the optimal strategies, these measures D_t^+ and D_t^- are positively correlated with the welfare of agents with $\Theta_i = t$. I illustrate what these measures look like for agents with different social types with the use of the running example.

Example - Consider again the setting in which high school students choose whether to enter a math competition or some mixed ability competition. Table 1 evaluates D_t^+ and D_t^- for agents with $\Theta_i = t$.

t	D_t^+	D_t^-
11	*Positive*	0
00	0	*Positive*
10	*Positive*	*Positive*
01	*Positive*	*Positive*

Table 1: The potential to improve decision making for each social type $\Theta_i = t$

Table 1 shows that Asian male students can only increase the likelihood of choosing to enter the math competition when $\alpha > \gamma$, while non-Asian female students can only increase the likelihood of choosing not to enter the math competition when $\alpha < \gamma$. Asian female students and non-Asian male students, on the other hand, have the ability to decrease the likelihood of making both types of mistake. These students are therefore on average more likely to choose their welfare-maximizing task than Asian male students and non-Asian female students.

This example sheds light on one of the main insights of the model, namely that we can divide the set of social types T into two different categories. These different categories play a key role in determining aggregate choice behavior, and are defined in Definition 4.

DEFINITION 4: *For a given social context* Π *we define two categories of social types* Θ*:*

- *A social type* Θ *is* ***mixed****, when* $\underline{\eta}_t < 1 < \overline{\eta}_t$
- *A social type* Θ *is* ***one-sided****, when* $\underline{\eta}_t > 1$ *or* $\overline{\eta}_t < 1$

Using this definition, we can generalize the results from the example. Specifically, an *intersectional lens* sheds light on the following asymmetries.

PROPOSITION 2 (Potential to Improve Decision Making): *Simultaneous asymmetry* $\pi_x^k \neq p_x^k$ *along the lines of the observable characteristics* θ^A *and* θ^B *leads to inequalities in the potential to improve decision making across the different social types* $\Theta = t$ *with* $t \in T$*. Specifically, agents with mixed social types have more potential to improve decision making than agents with one-sided social types, which leads to an on average higher expected utility for agents with mixed social types.*

Proposition 2 shows how the multidimensionality of social identity only reinforces the potential to improve decision making for agents with *mixed* social types, while it provides no extra benefits for agents with *one-sided* social types. Intersectionality in individual decision making therefore disadvantages agents with *one-sided* social types relative to agents with a *mixed* social type.

Example - Let us now aggregate agents with a *One-Dimensional lens* along the lines of gender[15]. We obtain the potential to improve decision making of male students by aggregating this potential of Asian and non-Asian male students. Asian male students only have $D_t^+ > 0$, while for non-Asian male students both measures are positive. We obtain the potential to improve decision making of female students in a similar manner. Non-Asian female students only have $D_t^- > 0$, while both measures are positive for Asian female students. This shows that, on average, male and female students have a similar potential to correct for the possible mistakes they can make. The fact that there are inequalities in decision making across different types of male and female students is therefore only visible with the use of an *Intersectional lens*. This section shows therefore how analyzing data with an *Intersectional lens* sheds light on additional insights, that are not visible when using a *One-Dimensional lens*.

Although the potential to improve decision making across male and female students is similar, there is an asymmetry. On average, there are more male students that are potentially able to correct for the mistake of choosing not to enter the math competition while $\alpha > \gamma$. On the other hand, there are on average more female students that are potentially able to correct for the mistake of choosing to enter this competition while $\alpha < \gamma$. This asymmetry will play an important role in determining aggregate choice behavior across social groups. An *Intersectional lens* shows how this asymmetry is driven by the behavior of agents with *one-sided* social types.

3.2.2 Aggregate Choice Behavior across Social Groups

The inequalities and asymmetries in the potential to improve decision making across social types affect aggregate choice behavior. Specifically, they do not only determine the number of agents with a certain social type $\Theta = t$ that chooses the *Competitive* task, but also their average success rate conditional on choosing this task.

Selection and Population Effects - The behavior implied by Proposition 1 induces probabilities $\Phi_{\alpha,t,\sigma_i^*,\Pi}$ with which agents choose the *Competitive* task. In the following example, I illustrate how differences in these probabilities across agents with different types (α, t) drive differences in choice behavior and average success rates across social groups.

[15] An analysis along the lines of being Asian leads to similar insights.

Example - Let $s_i \in \{0,1\}$, where $s_i = 1$ when $\alpha_i > \gamma$. Let $\hat{p}^{\sigma_i}(s_i, t_i)$ be the belief of an agent with (s_i, t_i) playing σ_i. Figure 2 shows the induced probabilities with which agents enter the math competition. All male students and all Asian students with $\alpha > \gamma$ focus on the social identity cue based on the representation of respectively male and Asian students among those that were successful in the previous cohort. They therefore enter the math competition for all realizations of $\hat{\alpha} > \frac{\gamma}{\eta_{k,1}}$ with $k \in \{A, B\}$. Their probability of entering the competition is represented by the top arrows. Students that are both Asian and male are indifferent between using π_1^A or π_1^B. Non-Asian female students with $\alpha > \gamma$ do not take their social identity cues into account, and enter the math competition for all realizations of $\hat{\alpha} > \gamma$. Similar for Asian male students with $\alpha < \gamma$. Their induced probabilities to enter the math competition are represented by the middle two arrows. Finally, both all non-Asian students and all female students with $\alpha < \gamma$ focus on the social identity cue based on the representation of respectively non-Asian and female students among those that were successful in the previous cohort. They therefore enter the math competition for all realizations of $\hat{\alpha} > \frac{\gamma}{\eta_{k,0}}$ with $k \in \{A, B\}$. Their probability of entering the competition is represented by the lower two arrows. Students that are both non-Asian and female are indifferent between using π_0^A and π_0^B.

We first analyze the aggregate choice behavior with a *One-Dimensional lens.* To illustrate the results, I aggregate the data along the lines of gender, but aggregating data along the lines of being Asian would lead to similar insights. Figure 2 shows that male students have on average a larger probability to enter the competition than female students. This is driven by the fact that more male students can increase the likelihood to enter the math competition when $\alpha > \gamma$, while more female students are able to increase the likelihood of not entering the math competition when $\alpha < \gamma$. Because the population is arbitrarily large, these probabilities translate into population fractions. Therefore, when male students are overrepresented among the successful students in the previous cohort, this induces more male students to choose to enter the competition than female students. At the same time, Figure 2 shows that female students enter the math competition for on average higher values of $\hat{\alpha}$ than male students. Because $\hat{\alpha}$ is unbiased, this implies that, conditional on entering the math competition, female students have on average a higher success rate than male students. Because α and the observable characteristics are independently distributed over the population, this selection effect does not reverse the order on the social identity cues π_1^k and π_0^k. The population and selection effects are generalized in Corollary 1.

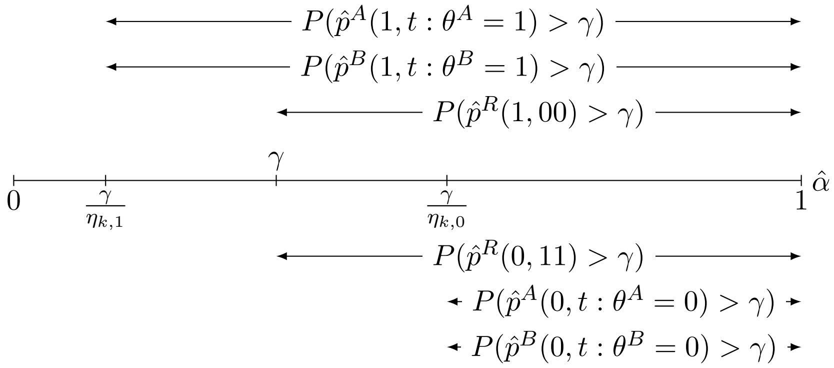

Figure 2: The induced probabilities for agents to choose $a = C$ implied by Proposition 1

COROLLARY 1.1 (One-Dimensional Lens): *For any observable characteristic θ^k such that $\pi_x^k > p_x^k$, we have population effect $\Phi_{\alpha,t:\theta^k=x,\sigma_i^*,\Pi} > \Phi_{\alpha,t:\theta^k=x',\sigma_i^*,\Pi}$ and a selection effect $E(\alpha|a = C, t : \theta^k = x) < E(\alpha|a = C, t : \theta^k = x')$. These effects are such that the order on π_x^k and $\pi_{x'}^k$ will not be reversed for any $k \in \{A, B\}$.*

An analysis with an *Intersectional lens* sheds light on what drives these *one-dimensional* selection and population effects. I illustrate the *intersectional* selection and population effects in the running example below.

Example - Figure 2 shows that Asian male students have on average the largest induced probability to enter the math competition. When they have $\alpha > \gamma$, they are able to increase the likelihood of entering the math competition, but when they have $\alpha < \gamma$, they cannot decrease the likelihood they mistakenly enter this competition. Non-Asian male and Asian female students have on average the second largest probability to enter the math competition. They can use their social identity cues to decrease the likelihood of making both types of mistake. Non-Asian female students have on average the smallest probability to enter the math competition, since they can only decrease the likelihood they mistakenly enter the competition when having $\alpha < \gamma$, but cannot increase the likelihood of entering the competition when having $\alpha > \gamma$. These different induced probabilities to enter the math competition induce *intersectional* population effects, such that

$$\Phi_{\alpha,11,\sigma_i^*,\Pi} > \Phi_{\alpha,10,\sigma_i^*,\Pi} = \Phi_{\alpha,01,\sigma_i^*,\Pi} > \Phi_{\alpha,00,\sigma_i^*,\Pi} \tag{12}$$

and consequently *intersectional* selection effects, such that

$$E(\alpha|a = C, t = 11) < E(\alpha|a = C, t \in \{10, 01\}) < E(\alpha|a = C, t = 00) \qquad (13)$$

These intersectional selection and population effects show there are no differences in choice behavior across non-Asian male students and Asian female students.

Let $\tilde{t}_x$ be the *one-sided* social type that has $\theta^k = x$ for all $k \in \{A, B\}$, while $\tilde{t}_{x'}$ is the *one-sided* social type that has $\theta^k = x'$ for all $k \in \{A, B\}$. Let t_{mixed} be any *mixed* social type. Then, the *intersectional* population and selection effects can be generalized as follows.

COROLLARY 1.2: *Let* Π *be such that* $\pi_x^k > p_x^k$ *for all* $k \in \{A, B\}$. *Then, we have an intersectional population effect, such that* $\Phi_{\alpha,\tilde{t}_x,\sigma_i^*,\Pi} > \Phi_{\alpha,t_{mixed},\sigma_i^*,\Pi} > \Phi_{\alpha,\tilde{t}_{x'},\sigma_i^*,\Pi}$, *and an intersectional selection effect, such that* $E(\alpha|a = C, \tilde{t}_x) < E(\alpha|a = C, t_{mixed}) < E(\alpha|a = C, \tilde{t}_{x'})$. *These effects are such that the order on* π_x^k *and* $\pi_{x'}^k$ *will not be reversed for any observable characteristic* $k \in \{A, B\}$.

Corollary 1.1 and 1.2 show how the individual behavior implied by Proposition 1 affects choice behavior at the aggregate level. Asymmetry $\pi_x^k \neq p_x^k$ along the lines of multiple observable characteristics θ^k influences the number of agents that chooses the *Competitive* task and their average success rate. Analyzing data with a *One-Dimensional lens* enables us to observe the differences across agents with a different value of each observable characteristic. An *Intersectional lens* shows nevertheless that these differences are really across social types, and that the *one-dimensional* population and selection effects are mainly driven by the behavior of agents with a *one-sided* social type.

How these *one-dimensional* and *intersectional* population and selection effects affect the possible equilibrium outcomes depends on whether they shrink or increase the differences in the induced social context π_x^k and π_t. I characterize the different possible equilibrium outcomes as follows.

DEFINITION 5 (Population Equilibrium): *In a* ***Neutral Regime*** *the allocation of individuals over tasks is symmetric across subgroups, and* $\pi_t = p_t$ *for all* $t \in T$. *In a* ***Non-Neutral Regime of Degree 2*** *the allocation of individuals over tasks is asymmetric along the lines of two observable characteristics, and* $\pi_x^k \neq p_x^k$ *for* $k = \{A, B\}$.

We can show that a *Neutral Regime* always exists. Imagine a social context Π such that $\pi_t = p_t$ for all $t \in T$, meaning no social type is relatively over- or underrepresented among the successful individuals. This implies that $\pi_x^k = p_x^k$ for all $k \in \{A, B\}$ and $x \in X$. Therefore, all strategies $\sigma \in \{A, B, R\}$ are equivalent, meaning that all individuals form beliefs in the same way, no matter their type (α, Θ). Therefore, there will be no differences in the induced choice behavior across social types, and $\tilde{\pi}_{k,x}(\sigma, \Pi) = p_x^k$ for all $k \in \{A, B\}$. To analyze when a *Non-Neutral Regime* exists, I define the following stability criterion. Let $\rho_k = \frac{\pi_x^k}{\pi_{x'}^k}$ be the fraction of the two social identity cues related to θ^k in a social context Π. Let $\tilde{\rho}_k = \frac{\tilde{\pi}_{k,x}(\sigma^*,\Pi)}{\tilde{\pi}_{k,x'}(\sigma^*,\Pi)}$ be the fraction of the social identity cues related to θ^k induced by the social context Π and the optimal strategies σ^*. Then, an equilibrium regime becomes unstable in dimension k when,

$$\frac{\partial \tilde{\rho}}{\partial \rho} > 1$$

In the following, I present an extreme example to illustrate how a *Neutral Regime* can become unstable and give rise to a stable *Non-Neutral Regime of Degree 2*.

Example - Consider the case in which the high school students have an extreme version of the response function η, such that,

$$\eta(\pi, p) = \begin{cases} +\infty & \text{if } \pi > p \\ 1 & \text{if } \pi = p \\ -\infty & \text{if } \pi < p \end{cases}$$

When $\pi_t = p_t$ for all $t \in T$, then all the strategies $\sigma \in \{A, B, R\}$ are equivalent and $\tilde{\pi}_{k,x}(\sigma, \Pi) = p_x^k$ for all $k \in \{A, B\}$. Now assume a small change in the social context, such that there are slightly more male and Asian students among the successful students in the previous cohort, such that $\pi_1^k > p_1^k$ for $k \in \{A, B\}$. Proposition 1 implies that, with an extreme response function η, this perturbation induces all male and Asian students with $\alpha > \gamma$ to enter the math competition, while no female students and non-Asian students with $\alpha < \gamma$ will enter the competition. Consequently, $\tilde{\pi}_1^k(\sigma, \Pi) > p_1^k$ for both $k \in \{A, B\}$, while $\tilde{\pi}_0^k(\sigma, \Pi) < p_0^k$, and the *Neutral Regime* becomes unstable.

To shed more light on how this perturbation exactly affects the induced social identity cues, let $S_1^A = \lim_{N\to\infty} \frac{N_1^A}{N}$ be the fraction of successful male students. When there is a perturbation in social context of the type mentioned above, all male students with $\alpha > \gamma$ will enter the math competition. Furthermore, no male non-Asian students with $\alpha < \gamma$ enter the competition, while male Asian students with $\alpha < \gamma$ only enter the competition when $\hat{\alpha} > \gamma$. Therefore,

$$S_1^A = p_1^A \int_{\alpha>\gamma} \alpha f(\alpha) d\alpha + p_1^A p_1^B \int_{\alpha<\gamma} \int_{\hat{\alpha}>\gamma} \alpha g_\alpha(\hat{\alpha}) f(\alpha) d\alpha d\hat{\alpha} \tag{14}$$

Let $S_0^A = \lim_{N\to\infty} \frac{N_0^A}{N}$ be the fraction of successful female students. The perturbation causes no female students with $\alpha < \gamma$ to enter the math competition. On the other hand, all Asian female students with $\alpha > \gamma$ will enter the competition, while non-Asian female students with $\alpha > \gamma$ only enter the competition when $\hat{\alpha} > \gamma$. Therefore,

$$S_0^A = p_0^A p_0^B \int_{\alpha>\gamma} \int_{\hat{\alpha}>\gamma} \alpha g_\alpha(\hat{\alpha}) f(\alpha) d\alpha d\hat{\alpha} + p_0^A p_1^B \int_{\alpha>\gamma} \alpha f(\alpha) d\alpha \tag{15}$$

Consequently, the social identity cue $\tilde{\pi}_1^A(\sigma, \Pi)$ induced by this extreme 'response function' is equal to,

$$\tilde{\pi}_1^A(\sigma, \Pi) = \frac{S_1^A}{S_1^A + S_0^A} \tag{16}$$

We can derive the induced social identity cue $\tilde{\pi}_1^B(\sigma, \Pi)$ in a similar manner. Because $S_1^k > S_0^k$ for both k, the induced social context will again be such that Asian and male students are relatively overrepresented among the successful students in the math competition, which will in turn induce again the behavior and statistics described above. The extreme 'response function' can therefore give rise to a stable *Non-Neutral Regime of degree 2*. Furthermore, since $S_0^k > 0$ for all k, it follows that $\tilde{\pi}_1^k(\sigma, \Pi) < 1$ for $k \in \{A, B\}$, meaning that there will be a positive number of non-Asian and female students among the successful individuals in the induced social context. Finally, $\tilde{\pi}_1^k(\sigma, \Pi)$ in increasing in the number of non-Asian female and Asian male students in the population, while it is decreasing in the number of Asian female students and non-Asian male students. This is not surprising, since we have seen that it is the behavior of agents with *one-sided* social types that mainly drives the differences in choice behavior across social groups.

An *Intersectional lens* provides additional insights on this matter. Let $S_t = \lim_{N\to\infty} \frac{N_t}{N}$ be the fraction of successful students with social type t. When $\pi_1^k > p_1^k$, all male Asian students with $\alpha > \gamma$ enter the math competition, while male Asian students with $\alpha < \gamma$ will only enter when $\hat{\alpha} > \gamma$. Therefore,

$$S_{11} = p_1^A p_1^B \int_{\alpha<\gamma} \int_{\hat{\alpha}>\gamma} \alpha g_\alpha(\hat{\alpha}) f(\alpha) d\alpha d\hat{\alpha} + p_1^A p_1^B \int_{\alpha>\gamma} \alpha f(\alpha) d\alpha \tag{17}$$

Similarly, no female non-Asian students with $\alpha < \gamma$ will enter the math competition, while female non-Asian students with $\alpha > \gamma$ will enter the competition when $\hat{\alpha} > \gamma$. This fraction is therefore bounded from below, such that

$$S_{00} = p_0^A p_0^B \int_{\alpha>\gamma} \int_{\hat{\alpha}>\gamma} \alpha g_\alpha(\hat{\alpha}) f(\alpha) d\alpha d\hat{\alpha} \tag{18}$$

Finally, all male non-Asian students and female Asian students with $\alpha > \gamma$ will enter the competition, while no students with these social types will enter the competition when $\alpha < \gamma$. Therefore,

$$S_{10} = p_1^A p_0^B \int_{\alpha>\gamma} \alpha f(\alpha) d\alpha \tag{19}$$

$$S_{01} = p_0^A p_1^B \int_{\alpha>\gamma} \alpha f(\alpha) d\alpha \tag{20}$$

Students with a *mixed* social type always choose their welfare-maximizing task when having an extreme 'response function', and their induced fractions S_t are equal. The differences between male (Asian) versus female (non-Asian) students in this extreme example are therefore completely driven by the fact that non-Asian female students cannot correct for the mistake of not entering the competition while $\alpha > \gamma$, while male Asian students cannot correct for the mistake of entering the competition while $\alpha < \gamma$. Equations (12) and (13) show that the behavior of agents with *mixed* social types only decreases the difference between S_1^k and S_0^k.

Proposition 3 generalizes these insights and shows how the *intersectional* population and selection effects determine the order on the π_t that can exist in a *Non-Neutral Regime of Degree 2.*

PROPOSITION 3: *There exists a **Non-Neutral Regime of degree 2** in which WLOG $\pi_x^k > p_x^k$ for all $k \in \{A, B\}$. In such a regime, any possible order on the social identity cues must be such that,*

$$\pi_{\tilde{t}_{x'}} = \min_{k,x} \pi_x^k \qquad \pi_{\tilde{t}_x} = \max_{k,x} \pi_x^k$$

where $\tilde{t}_x$ is such that $\theta^k = x$, while $\tilde{t}_{x'}$ is such that $\theta^k = x'$ for all $k \in \{A, B\}$.

Proposition 3 shows that the differences in choice behavior induced by social context can give rise to persistent differences in the representation of social groups among the successful individuals at the *Competitive* task. Agents with a social type such that they belong to the socially more successful group according to both observable characteristics will be most overrepresented among the successful individuals. This is driven by the fact that these agents are most likely to mistakenly choose the *Competitive task* while $\alpha < \gamma$. Agents with a social type such that they belong to the socially less successful group according to both observable characteristics will be most underrepresented among the successful individuals. This is driven by the fact that these agents are most likely to mistakenly choose the *Non-Competitive* task, even though $\alpha > \gamma$.

To summarize, when $\pi_x^k \neq p_x^k$ along the lines of multiple observable characteristics, this only reinforces the potential to improve decision making of agents with *mixed* social types. These agents can use their two observable characteristics to decrease the likelihood of mistakenly choosing the *Competitive* task while $\alpha < \gamma$, and of mistakenly choosing the *Non-Competitive* task while $\alpha > \gamma$. Agents with *one-sided* social types can only potentially correct for one of these two mistakes. There is nevertheless an asymmetry in the potential to improve decision making across agents with *one-sided* social types, where they have the ability to potentially correct for a different type of mistake. The differences and asymmetry in the potential to improve decision making induce differences in choice behavior. An *Intersectional lens* sheds light on the fact that differences in *one-dimensional* statistics are mainly driven by the behavior of agents with *one-sided* social types. Finally, these differences in choice behavior can be persistent. Specifically, the differences in the potential to improve decision making between *mixed* and *one-sided* social types, and the asymmetry in this potential across *one-sided* social types give rise to persistent differences in the representation of these social types among the successful individuals at the *Competitive* task.

3.3 Adding Two-Dimensional Social Identity Cues

To simplify the exposition of the model, I assumed agents could only use *one-dimensional* social identity cues. In this section, I show what the effects are of adding the option of using the *two-dimensional* social identity cues π_t in belief formation. This changes the strategy set to $\sigma_i \in \{A, B, F, R\}$, where F refers to the strategy in which agents use the *two-dimensional* social identity cue derived from their full social type Θ_i. When $\sigma_i = F$, the corresponding belief $\hat{p}_i^F = \eta_t \hat{\alpha}_i$, where, $\eta_t = \eta(\pi_t, p_t)$. The function $\eta(\pi_t, p_t)$ can be different from $\eta(\pi_x^k, p_x^k)$. People could for example react stronger to *two-dimensional* social identity cues than *one-dimensional* social identity cues. Consequently, let

$$\overline{\overline{\eta}}_t = \max_{t,k\in\{A,B\}} \eta_{k,x}, \eta_t \text{ and } \underline{\underline{\eta}}_t = \min_{t,k\in\{A,B\}} \eta_{k,x}, \eta_t$$

Similarly, define

$$\overline{\kappa}_t = \operatorname*{argmax}_{t,k\in\{A,B\}} \eta_{k,x}, \eta_t \text{ and } \underline{\kappa}_t = \operatorname*{argmin}_{t,k\in\{A,B\}} \eta_{k,x}, \eta_t$$

This shows that $\overline{\overline{\eta}}_t \geq \overline{\eta}_t$, while $\underline{\underline{\eta}}_t \leq \underline{\eta}_t$. Adding the strategy F provides the following Corollary to Proposition 1.

COROLLARY 2: *Consider an agents with social type $\Theta_i = t$. When this agents has $\alpha > \gamma$, $\sigma_i^* = \overline{\kappa}_t$ if and only if $\overline{\overline{\eta}}_t > 1$. Otherwise, $\sigma_i^* = R$. When agents have $\alpha < \gamma$, then $\sigma_i^* = \underline{\kappa}_t$ if and only if $\underline{\underline{\eta}}_t < 1$. Otherwise, $\sigma_i^* = R$.*

Corollary 2 shows that agents will only use the option $\sigma_i = F$, when this leads to a bias of their noisy perception $\hat{\alpha}$ in the direction of their welfare maximizing task, and when η_t provides a stronger bias in this direction than $\eta_{k,x}$ for any k. The following example illustrates how this extra strategy can affect the outcomes at the aggregate level.

Example - In the previous section, I showed how female non-Asian students with $\alpha < \gamma$ were indifferent between choosing to use π_0^A or π_0^B. If Corollary 2 implies they now use the cue π_t, then it must be that $\eta_{00} < \eta_{k,0}$ for $k \in \{A, B\}$. The opposite applies to male Asian students, who were indifferent between choosing to use π_1^A or π_1^B. When $\eta_{00} < \eta_{k,0}$ and $\eta_{11} > \eta_{k,1}$ for $k \in \{A, B\}$, both female non-Asian and male Asian students choose $\sigma_i = F$. This increases the strength of the *intersectional* population and selection effects across female non-Asian and male Asian students, which consequently increases the strength of the *one-dimensional* population and selection effects across male (Asian) and female (non-Asian) students. This insight is generalized in Corollary 3

COROLLARY 3: *The introduction of the strategy $\sigma_i = F$ does not invalidate the existence of a Non-Neutral Regime of Degree 2.*

Example - In the previous example, the strategy $\sigma_i = F$ reinforces the *one-dimensional* population and selection effects. This is not always the case. It could be that η_t induces a much stronger bias for a similar pair (π, p) than $\eta_{k,x}$. In this case, it could happen that female Asian and male non-Asian students also end up choosing $\sigma_i = F$. When this is the case, agents with *mixed* social types are still able to decrease the likelihood of making both types of mistakes. Because η_{01} and η_{10} are either above or below one, the use of the extra option t can only improve the ability to potentially correct for one of the two possible mistakes. This induces two possible settings.

In the first setting, both Asian and non-Asian female students are relatively underrepresented, while both Asian and non-Asian female students are relatively underrepresented. Asian female students will now only choose $\sigma_i = F$ when $\alpha < \gamma$. This increases the ability for Asian female students to potentially correct for the mistake of choosing to enter the math competition while $\alpha < \gamma$ compared to their ability to potentially correct for the mistake of not entering the competition while $\alpha > \gamma$ following from $\sigma_i = k$. Non-Asian male students will only choose $\sigma_i = F$ when $\alpha > \gamma$. The option F in this setting increases their ability to potentially correct for not entering the math competition while $\alpha > \gamma$ compared to their ability to potentially correct for entering the math competition while $\alpha < \gamma$ following from their strategy $\sigma_i = k$. When we evaluate these changes with a *one-dimensional lens*, it follows that the average probability to enter the math competition decreases for female and Asian students, while it increases for male and Non-Asian students. Therefore, the behavior of agents with *mixed* social types following the introduction of the strategy F increases the strength of the population and selection effects along the lines of gender, while it decreases the strength of the population and selection effect along the lines of being Asian. The opposite happens when Asian female students are relatively overrepresented, while non-Asian female students are relatively overrepresented, while vice versa for male students. The net effect of the introduction of the strategy F depends on total effect of this strategy on the behavior of agents with both *mixed* and *one-sided* social types

To summarize, the introduction of the strategy $\sigma_i = F$ can further reinforce the asymmetry in decision making across agents with a *one-sided* type. This reinforces the induced population and selection effects. The strategy $\sigma_i = F$ can nevertheless also introduce asymmetries in the potential to improve decision making across agents with *mixed* social types. This type of asymmetry affects the *one-dimensional* population and selection effects of the two dimensions of social identity in opposite directions, reinforcing these effects for one dimension of social identity, while decreasing these effects for the other dimension of social identity.

3.4 Discussion

Most assumptions of the model are discussed in Liqui Lung (2022). To simplify the discussion of the multidimensionality of social identity, I made one additional assumption, namely that the individual observable characteristics are independently distributed over the population. Although this is a reasonable assumption to make in many cases, there are also observable characteristics that are correlated. When this is the case,

$$p_t \neq p_x^A p_x^B$$

In most cases, two social identities that both imply belonging to the socially less or more successful group are positively correlated. For example, belonging to an underrepresented minority group is often positively correlated with belonging to a lower income class. When this type of correlation exists, there will be relatively more agents with a *one-sided* social type than a *mixed* social type. Because agents with *one-sided* social types mainly drive the *one-dimensional* population and selection effects, this type of correlation leads to an increase in the strength of these effects. This induces larger differences between the overrepresentation of agents with the socially more successful *one-sided* social type, and the underrepresentation of agents with the socially less successful *one-sided* social type among the successful individuals in the *Competitive* task. When one social identity that implies belonging to the socially less successful group and another that implies belonging to the socially more successful group are positively correlated, this increases the fraction of agents with a *mixed* social type relative to the number of agents with a *one-sided* social type. This decreases the strength of the *one-dimensional* population and selection effects, and decreases the differences in the representation of different social types among the successful individuals.

4 Social Constraints

In the previous sections, I assumed that agents could freely choose to ignore one dimension of their social identity. In reality, this ability may be limited because of socially imposed constraints, such as the existence of stereotypes, narratives, stigmatization or the inability to identify with certain social types. The framework developed in this paper allows us to study the effects of these social constraints in terms of strategy restrictions. I group these restrictions into two types. The first type is such that agents are not able to ignore one dimension of their social type, while they are free to ignore the other dimension. I summarize constraints of this type under the name *Stigmatization*. The second type of restriction is such that agents are unable to ignore any dimension of their social type, and can only use *two-dimensional* social identity cues in belief formation. I refer to this type of constraint as *Type-Specific Social Identification*.

4.1 Stigmatization

The restriction that some agents are not able to ignore one of the dimensions of their social identity aims to capture the consequences of issues such as stigmatization, stereotypes, narratives and the fact that one's social identity is a composite view of the view one has of oneself as well as the views held by others about one's identity (Nagel, 1994). In the following, I introduce this restriction in the model in two ways. In the first version, an entire observable characteristic is stigmatized, which implies that agents cannot ignore this characteristic. I first analyze a *Two-Strategy Model*, in which agents cannot use the social identity cues π_t. Consequently, I introduce the *Three-Strategy Model*, in which agents can use both the cues π_x^k and π_t. In the second version, only one value of an observable characteristic is stigmatized, meaning that only those agents with that specific value of the characteristic cannot ignore the characteristic, while agents with another value can. I call this model the *Asymmetric Model*.

4.1.1 The Two-Strategy Model

In the *Two-Strategy Model*, agents can only use *one-dimensional* social identity cues. Assume gender is the stigmatized dimension of social identity. The strategy set is therefore reduced to $\sigma_i \in \{A, R\}$. The optimal strategies for each social group are presented in Figure (3). This figure shows there is no difference anymore in the potential to improve decision making between *mixed* and *one-sided* social types. Therefore, only agents with *mixed* types are disadvantaged by stigmatization in this model, while agents with *one-sided* types are not effected at all. Specifically, Asian female students lose the ability to decrease the likelihood they choose not to enter the math competition when $\alpha > \gamma$, while non-Asian male students lose the ability to decrease the likelihood of choosing to enter the math competition while $\alpha < \gamma$. Consequently, agents with *mixed* and *one-sided* social types now have the exact same potential to improve decision making using social identity cues.

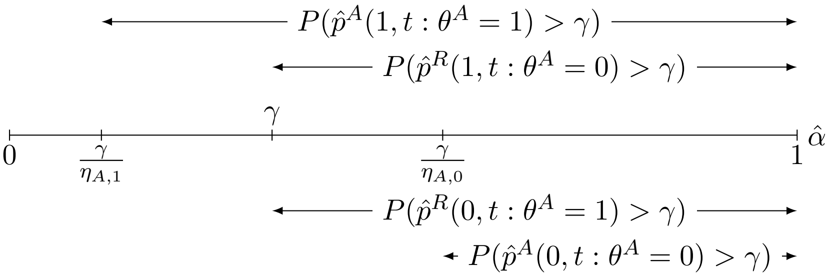

Figure 3: The optimal strategies for each social type in a Two-Strategy model.

Furthermore, agents cannot use the dimension of being Asian in decision making. Because the characteristics θ^k and α are independently distributed over the population, there can be no differences in choice behavior between Asian and non-Asian students. Furthermore, the differences in choice behavior between male and female students are now driven by the behavior of agents with both one-sided social types and *mixed* social types. Therefore, the *Two-Strategy* model induces larger population and selection effects in the dimension of gender than a model without stigmatization[16].

4.1.2 The Three-Strategy Model

In the *Three-Strategy Model*, agents can use both *one-dimensional* and *two-dimensional* social identity cues. When gender is stigmatized, the strategy set becomes $\sigma_i \in \{A, F, R\}$. Like in Section 3.3, there are two different settings. In the first setting, both Asian and non-Asian male students are relatively overrepresented among the successful students in the previous cohort, while both Asian and non-Asian female students are relatively underrepresented. In this case, Asian female students with $\alpha > \gamma$ and non-Asian male students with $\alpha < \gamma$ cannot use $\sigma_i = F$ to improve decision making. Consequently, the *Three-Strategy* model has similar implications as the *Two-Strategy* model, where *mixed*

[16] Like in Section 3.3, there is an asymmetry in the potential to improve decision making across *mixed* social types. Since agents with *mixed* social types now act exactly as agents with *one-sided* social types, this model presents an extreme case. Consequently, the population and selection effects in the dimension of gender are reinforced, while the population and selection effects in the dimension of being Asian completely disappear.

social types lose their ability to decrease the likelihood of making one type of mistake. When agents with *one-sided* social types find it optimal to use π_t instead of π_x^A, then the optimal strategy of agents with *one-sided* social types induces a larger bias than the optimal strategy of agents with *mixed* social types. This creates a difference in the potential to improve decision making between *mixed* social types and *one-sided* social types, which induces differences in choice behavior across Asian and non-Asian students. The availability of π_t can therefore induce population and selection effects in dimensions of social identity that are not stigmatized. The population and selection effects in the dimension of gender are potentially larger than in a model without stigmatization, since they will be driven by both the behavior of agents with *mixed* and *one-sided* social types. The population and selection effects in the dimension of being Asian are smaller, since differences in choice behavior between Asian and non-Asian students are only driven by the fact that non-Asian female students can slightly better correct for the mistake of entering the competition while $\alpha < \gamma$ than Asian female students, while Asian male students can slightly better correct for the mistake of not entering the competition while $\alpha > \gamma$ than non-Asian male students.

In the second setting, Asian male students are relatively underrepresented, but non-Asian male students are relatively underrepresented. Similarly, non-Asian female students are relatively underrepresented, while Asian female students are relatively overrepresented.

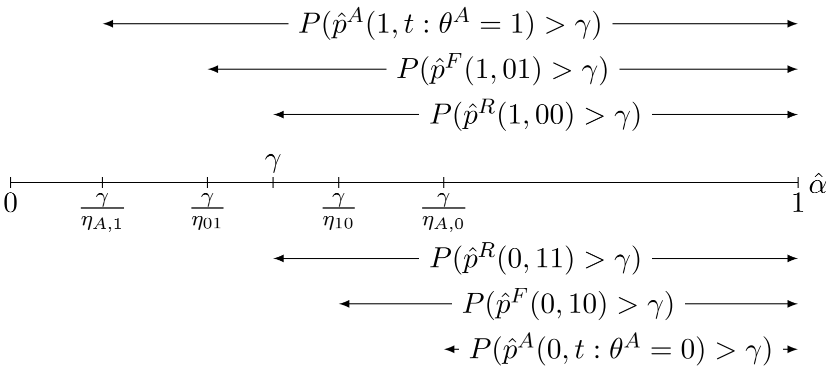

Figure 4: The optimal strategies for each social type in a Three-Strategy model, where we assume for simplicity that $\frac{\gamma}{\eta_{A,0}} = \frac{\gamma}{\eta_{00}}$ and $\frac{\gamma}{\eta_{A,1}} = \frac{\gamma}{\eta_{11}}$

Figure 4 shows that non-Asian male and Asian female students maintain their ability to potentially correct for both types of mistakes. In this setting, the availability of the cues π_t can therefore enable agents with *mixed* social types to escape the negative effects of the stigmatization of gender. There will be a difference in the potential to improve decision making across agents with *mixed* social types, and the results are similar to those presented in section 3.3.

4.1.3 The Asymmetric Model

In the *Asymmetric* model, only one value of an observable characteristic is stigmatized. Assume that in our example being female is stigmatized. This means that female students cannot ignore being female when forming beliefs, which implies their strategy set is $\sigma_i \in \{A, F, R\}$. Male students, on the other hand, can use the complete strategy set $\sigma_i \in \{A, B, F, R\}$. First, consider the setting in which Asian male students are relatively overrepresented, but non-Asian male students are relatively underrepresented, while vice versa for female students. In this setting, Asian female students can potentially escape the negative effects of stigmatization by using the cue π_t instead of π_1^B. If it is optimal to choose $\sigma_i = F$ for all agents with a *mixed* social type, stigmatization does not affect the results at the aggregate level. If it is optimal to choose $\sigma_i = k$ for all agents with a *mixed* social type, then the stigmatization of being female negatively affects Asian female students with $\alpha > \gamma$. They have slightly less potential to decrease the likelihood of not entering the math competition than non-Asian male students. This creates a difference, but no asymmetry, in the potential to improve decision making between Asian female and non-Asian male students. Therefore, the potential of male and non-Asian students to improve decision making slightly increases, while it slightly decreases for female and Asian students. As a result, the population and selection effects in the dimension of gender will be reinforced, while the strength of the population and selection effects in the dimension of being Asian decreases.

These effects are much stronger when both Asian and non-Asian male students are relatively overrepresented among the successful students in the previous cohort, while both Asian and non-Asian female students are relatively underrepresented. Figure 5 shows how this affects the optimal strategies for agents with different social types. In this setting, Asian female students are no longer able to potentially correct for the mistake of choosing not to enter the competition while $\alpha > \gamma$. This reinforces the difference in the potential to improve decision making between non-Asian male students and Asian female students, while the potential of non-Asian female and Asian male students is left unchanged. In this model, stigmatization increases therefore the population and selection effects in the dimension of the stigmatized dimension of social identity, while it decreases these effects in the dimension of the non-stigmatized dimension of social identity.

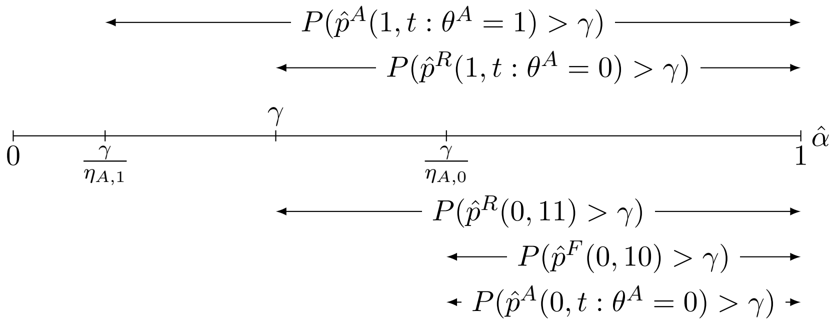

Figure 5: The optimal strategies for each social type in an Asymmetric model. We assume for simplicity that $\frac{\gamma}{\eta_{A,0}} = \frac{\gamma}{\eta_{00}}$ and $\frac{\gamma}{\eta_{A,1}} = \frac{\gamma}{\eta_{11}}$

4.2 Type-Specific Social Identification

The second type of restriction implies that agents cannot ignore any dimension of their social type in belief formation. This type of restriction aims to capture the phenomenon that agents with certain social types have very different experiences in society, which makes them unable to identify with individuals from another social type. This type of restriction reduces the strategy set to $\sigma_i \in \{F, R\}$. The first insight this restriction provides is it that, when we eliminate the option to use *one-dimensional* social identity cues, all social types only have the ability to decrease the likelihood of making one type of mistake. There are again two different settings that can arise. First, both Asian and non-Asian male students may be relatively overrepresented among the successful students in the previous cohort, while both Asian and non-Asian female students are relatively underrepresented. In this setting, both Asian and non-Asian male students have the ability to increase the likelihood of entering the math competition when $\alpha > \gamma$, while both Asian and non-Asian women have the ability to decrease the likelihood of entering the math competition when $\alpha < \gamma$. This induces population and selection effects in the dimension of gender that are driven by the behavior of agents with both *one-sided* and *mixed* social types. The story is different when we evaluate asymmetry along the dimension of being Asian. Asian male students have the ability to increase the likelihood of entering the competition while $\alpha > \gamma$, while Asian female students have the ability to decrease the likelihood of entering the competition when $\alpha < \gamma$. For non-Asian male students and non-Asian female students, this is exactly the opposite. To give more insight into what happens to the induced number of successful Asian and non-Asian students, let us go back to the extreme 'response function'. In this case, all male Asian students with $\alpha > \gamma$ and those with $\alpha < \gamma$ and $\hat{\alpha} > \gamma$ will enter the math competition, while only the female Asian students with $\alpha > \gamma$ and $\hat{\alpha} > \gamma$ enter the competition. The fraction of Asian students among all successful individuals induced in this model is given by,

$$S_1^B = p_1^A p_1^B \int_{\alpha>\gamma} \alpha f(\alpha) d\alpha + p_0^A p_1^B \int_{\alpha>\gamma} \int_{\hat{\alpha}>\gamma} \alpha g_\alpha(\hat{\alpha}) f(\alpha) d\alpha d\hat{\alpha} + p_1^A p_1^B \int_{\alpha<\gamma} \int_{\hat{\alpha}>\gamma} \alpha g_\alpha(\hat{\alpha}) f(\alpha) d\alpha d\hat{\alpha} \tag{21}$$

Similarly, all male non-Asian students with $\alpha > \gamma$ and those with $\alpha < \gamma$ and $\hat{\alpha} > \gamma$ will enter the math competition, while only the female non-Asian students with $\alpha > \gamma$ and $\hat{\alpha} > \gamma$ enter the competition. The fraction of non-Asian students among all successful individuals is equal to,

$$S_0^B = p_1^A p_0^B \int_{\alpha>\gamma} \alpha f(\alpha) d\alpha + p_0^A p_0^B \int_{\alpha>\gamma} \int_{\hat{\alpha}>\gamma} \alpha g_\alpha(\hat{\alpha}) f(\alpha) d\alpha d\hat{\alpha} + p_1^A p_0^B \int_{\alpha<\gamma} \int_{\hat{\alpha}>\gamma} \alpha g_\alpha(\hat{\alpha}) f(\alpha) d\alpha d\hat{\alpha} \tag{22}$$

When $p_1^k = p_0^k$ for all $k \in \{A, B\}$, these two fractions are equal. Furthermore, an increase in the fraction of the population with *one-sided* social types increases S_1^B relative to S_0^B. Without an extreme 'response function', there will only be differences in S_1^B and S_0^B when the social context is such that $\pi_{11} \neq \pi_{10}$ and $\pi_{00} \neq \pi_{01}$.

The second setting that can arise is when Asian male students are relatively underrepresented, but non-Asian male students are relatively underrepresented. Similarly, non-Asian female students are relatively underrepresented, while Asian female students are relatively overrepresented. In this case, the exact opposite happens. Now, all Asian students can potentially correct for not entering the math competition while $\alpha > \gamma$, while all non-Asian students can potentiallly correct for entering the math competition while $\alpha > \gamma$. Differences in choice behavior along the lines of being Asian are therefore driven by the behavior of agents with both *mixed* and *one-sided* types. Along the lines of gender, now Asian male (female) students and non-Asian male (female) students have the ability to potentially correct for a different type of mistake. This setting produces therefore the exact opposite results.

4.3 Discussion

An analysis of the effects of social constraints provides two main insights. First, any restrictions on the strategy set of agents mainly affect the choice behavior of agents with *mixed* social types. Specifically, social constraints can limit their potential to improve decision making by taking away the ability to decrease the likelihood of making a certain type of mistake. The *Two-Strategy* model shows that, without the option to use *two-dimensional* social identity cues, stigmatization takes away all the advantages *mixed* social types had relative to *one-sided* social types. The *Three-Strategy* and *Asymmetric* model show that the availability of the *two-dimensional* social identity cues π_t can under certain circumstance give *mixed* social types the ability to escape the negative effects of stigmatization. A comparison of stigmatization in the different model sheds therefore light on the importance of the availability of *two-dimensional* statistics. Finally, a model with *Type-Specific Social Identification* shows that it are also the agents with *mixed* social types that are disadvantaged when we take away the option to use *one-dimensional* social identity cues. Especially for these agents, it is therefore important that both *one-dimensional* and *two-dimensional* social identity cues are available, and that agents are able to identify with multiple subgroups.

Secondly, stigmatization increases the strenght of the population and selection effects in the dimension of the stigmatized dimension of social identity. A comparison of the different models of stigmatization shows how the difference in the potential to improve decision making between *mixed* and *one-sided* social types plays a role in what happens along the lines of the other dimension of social identity. In the *Two-Strategy* model, we saw that, when both social types behave the same, there can only be asymmetry along the lines of the stigmatized observable characteristic. The *Three-Strategy* model shows that, when there is a small difference in the potential to improve decision making between *one-sided* and *mixed* social types, this is sufficient to induce small population and selection effects in the non-stigmatized dimension of social identity. The *Asymmetric* model shows that, when stigmatization creates a difference in the potential to improve decision making among agents with different *mixed* social types, this induces larger population and selection effects along the lines of the non-stigmatized observable characteristic, but they are still smaller than along the lines of the stigmatized social identity. Finally, in a model with *Type-Specific Identification*, both agents with *mixed* and *one-sided* social types can only decrease the likelihood of making one type of mistake. When agents with a *mixed* and *one-sided* social type can potentially correct of the same type of mistake, this increases the population and selection effects along the lines of this dimension of social identity. This nevertheless implies that, in the dimension of the other observable characteristics, agents with a *mixed* and *one-sided* social type can each decrease the likelihood of making a different type of mistake. The population and selection effects along the lines of this observable characteristic will therefore be small. These differences in the induced population and selection effects could be used to empirically discriminate between the different models and identify the type of social constraint in a setting.

5 Affirmative Action Policy and Intersectionality

In this section, I use the framework to analyze the effects of affirmative action policy. I show how an intersectional view and knowledge of the social constraints agents face are crucial for the development of these policies. Without it, policies may not have the desired outcomes, may not affect the agents we target, or may even transfer inequalities from one dimension of social identity to another. I use an example that is often studied in the literature[17], namely the use of quotas to increase the political participation of women and minorities. Specifically, I analyze on how the outcomes of one round of affirmative action policy affect the decision to run for office in future generations[18].

Let gender be denoted by $\theta^A \in \{\text{M,F}\}$, representing respectively male and female agents. Let $\theta^B \in \{\text{N-URM,URM}\}$ represent whether an individual belongs to an underrepresented minority (URM) or not (N-URM). I simplify notation by writing $\eta_x \equiv \eta_{k,x}$. Assume that, traditionally, female candidates and candidates belonging to an URM are underrepresented in the political institution we consider. Specifically, before affirmative action policy, social context was such that male and N-URM candidates were overrepresented to the same degree, such that $\eta_{\text{M}} = \eta_{\text{N-URM}}$, while female and URM candidates were underrepresented to the same degree, such that $\eta_{\text{F}} = \eta_{\text{N-URM}}$. Furthermore, assume that male N-URM candidates are most overrepresented, while female URM candidates are most underrepresented, and that $\eta_{\text{M,N-URM}} > \eta_{\text{M,URM}} > \eta_{\text{F,N-URM}} > \eta_{\text{F,URM}}$. Consequently, I analyze the effects of a social context that arises after having imposed a quota for female candidates on the decision of agents in the next generation to run for office. The quota could have affected social context in three different ways.

[17]See for example Cassan and Vandewalle (2021), Beaman et al. (2012), Hughes (2011), Karekurve-Ramachandra and Lee (2020), Folke et al. (2015) and Tan (2014)

[18]The probability to get elected can be correlated with belonging to certain social groups. For simplicity, I do not take into consideration that agents may account for this when choosing whether to run for office. This would nevertheless only further reinforce the effects I present in this section.

First, the quota may have affected the representation of URM and N-URM female candidates equally. Let us start with the assumption that agents can only use *one-dimensional* social identity cues. The quota increased the representation of female candidates relative to male candidates. This leads to an increase in η_{F} and a decrease in η_{M} for agents in the next generation. This change does not affect the decision making of male N-URM agents, who will use $\pi_{\text{N-URM}}$ when $\alpha > \gamma$ and cannot improve decision making when $\alpha < \gamma$. Similarly, it will not affect the decision making of female URM agents, who will use π_{URM} when $\alpha < \gamma$ and cannot improve decision making when $\alpha > \gamma$. It will nevertheless affect the decision making of agents with *mixed* social types. Specifically, following the decrease in η_{M}, male URM agents are now less likely to run for office when $\alpha > \gamma$. Similarly, following the increase in η_{F}, female N-URM agents are now more likely to run for office when $\alpha < \gamma$. When the quota erases all differences between male and female agents, these *mixed* social types lose their ability to correct for the respective mistakes. The quota therefore indeed decreases differences in choice behavior between female and male agents, but only through a decrease in the number of URM male agents, and an increase in the number of N-URM female agents that run for office. Consequently, it also increases differences in decision making between N-URM and URM agents, and transfers inequalities from the dimension of gender to the dimension of belonging to an URM. The availability of *two-dimensional* social identity cues can potentially reduce this spillover effect on the decision making of URM versus N-URM agents[19]. When gender is stigmatized, the quota is equally effective, but the spillover effects of the gender quota on the differences in decision making between N-URM and URM agents are smaller[20]. On the other hand, when the URM dimension of identity is stigmatized, a gender quota has very little effect. It can only lead to a small decrease in the differences in decision making between male and female agents, when $\eta_{\text{M,N-URM}} > \eta_{\text{N-URM}}$, while $\eta_{\mathrm{F,URM}} < \eta_{\mathrm{URM}}$, and male N-URM and female URM agents use their *two-dimensional* cues in decision making. On the other hand, when both dimensions of social identity are stigmatized, agents can only identify with others with the same social type. In this case, a quota that affects all female agents equally is very effective and decreases the differences in decision making in both dimensions of social identity simultaneously.

[19] When agents can also use *two-dimensional* social identity cues, and $\eta_{\text{M,N-URM}} > \eta_{\text{N-URM}}$, while $\eta_{\mathrm{F,URM}} < \eta_{\mathrm{URM}}$, male N-URM and female URM agents will use their *two-dimensional* cues in decision making. Because the quota increases $\eta_{\mathrm{F,URM}}$ and reduces $\eta_{\text{M,N-URM}}$, this reduces the population and selection effects in the dimension of belonging to an URM.

[20] Stigmatization mainly affects the decision making of agents with a *mixed* social type. In this case, male URM agents with $\alpha < \gamma$ cannot use the cue based on URM agents to improve decision making, and are more likely to run for office. Similarly, female N-URM agents with $\alpha > \gamma$ cannot use the cue based on N-URM agents to improve decision making, and are less likely to run for office. This dampens the spillover effect of the quota on the differences in decision making between N-URM and URM agents.

In practice, the quota may nevertheless affect the representation of N-URM female agents more than the representation of URM female agents[21]. Let us analyze the extreme case in which the quota only enhances the representation of N-URM female candidates. This leads again to an increase in η_{F} and a decrease in η_{M}. Therefore, female N-URM agents with $\alpha < \gamma$ are again more likely to run for office, while male URM agents with $\alpha > \gamma$ are less likely to do so. The quota now nevertheless also leads to an increase in $\eta_{\text{F,N-URM}}$, which induces a decrease in all other *two-dimensional* social identity cues. When $\eta_{\text{F,URM}} < \eta_{\text{URM}}$, the decrease in $\eta_{\text{F,URM}}$ causes female URM agents with $\alpha < \gamma$ to be less likely to run for office. While the quota enhances the representation of female N-URM agents among those who run for office, it therefore only further decreases the representation of female URM agents. This effect is even stronger, when belonging to an URM or both dimensions of social identity are stigmatized.

Finally, in the previous sections, we saw that differences in decision making along the lines of a single dimension of social identity are mainly driven by agents with *one-sided* social types. We could therefore consider a quota that specifically targets the representation of these agents. Such a quota would lead to a simultaneous increase in η_{F}, η_{URM} and $\eta_{\text{F,URM}}$, while it leads to a decrease in η_{M}, $\eta_{\text{N-URM}}$ and $\eta_{\text{M,N-URM}}$. This affects the decision making of agents with both *mixed* and *one-sided* social types, and leads to a simultaneous decrease in differences in decision making across both male versus female agents, and N-URM versus URM agents. Such a quota avoids therefore that differences in decision making along the lines of one dimension of social identity are transferred to another dimension. Furthermore, targeting agents with *one-sided* social types increases the effectiveness of a quota when one dimension of social identity is stigmatized[22]. When both dimensions of social identity are stigmatized, and social identification can only be type-specific, this type of quota is nevertheless not the optimal choice. When $\eta_{\text{F,URM}}$ increases, this automatically leads to a decrease in the relative representation of all other social types. This will therefore decrease the probability of all male agents to run for office, but it also decreases the probability of N-URM female agents to run for office. A policy that targets all female agents equally will therefore be more effective in this setting.

[21] See Crenshaw (1991) and Yuval-Davis (2006) for a discussion on how policies that aim to enhance the opportunities of women disproportionally affect white women.

[22] A quota targeting agents with *one-sided* social types has a similar effect on the social identity cues related to both dimensions of social identity. This type of quota therefore affects the behavior of agents with *mixed* social types, no matter which dimension of social identity is stigmatized.

To summarize, a gender quota has the risk of increasing inequalities along the lines of another dimensions of social identity. Furthermore, the effectiveness of this quota can be limited when other dimensions of social identity are stigmatized. When only one or none of the dimensions of social identity are stigmatized, a quota targeting agents with *one-sided* social types is most effective. When both dimensions of social identity are stigmatized, a gender quota is most effective when it targets the representation of URM and N-URM female agents equally.

Finally, when agents are perfectly able to learn their optimal strategies, the effects of a quota in this model are probably not achieved in the way policy makers would desire. Specifically, the changes are obtained through a loss of male candidates with $\alpha > \gamma$, and a gain in female candidates with $\alpha < \gamma$ that run for office. This sheds light on how affirmative action policy can potentially hurt the decision making of both female and male agents, which decreases aggregate welfare. This suggests that, in the absence of discrimination, an increase in the number of female and/or minority agents with $\alpha > \gamma$ that choose to run for office is possibly better obtained with the use of informational policies. Examples are policies that change the type of statistics that people process[23] or policies that make other dimensions of social identity salient that female and minority candidates with $\alpha > \gamma$ can use to increase the likelihood they will run for office.

6 Conclusion

This paper shows how intersectionality plays a role in individual decision making, and how analyzing data with an *intersectional lens* sheds light on inequalities that are invisible when only using a *one-dimensional* lens. An intuitive way to make a distinction between social groups in the development of policy would be to separate individuals along the lines of whether they belong to the socially more or less successful group in society. The key insight of this paper is nevertheless that, although there is an asymmetry, there are no differences between these two subgroups in their potential to improve decision making. Instead, the relevant distinction to make is between the following two social types: *mixed* social types, who are belong the socially more successful group along the lines of one dimension of their social identity, while they belong to the socially less successful group along the lines of their other dimension of social identity, and *one-sided* social types, who belong to either the socially more or less successful group along the lines of all the dimensions of their social identity.

[23]See Liqui Lung (2022) for a discussion on the effects of the structure of information on behavior.

Agents with *one-sided* social types are relatively disadvantaged, no matter whether they belong to the socially more or less successful group in society. Those that belong to the former group are disadvantaged because they are relatively more likely to choose a *Competitive* task when this is not their welfare maximizing choice. Those that belong to the socially less successful group are disadvantaged because they are relatively more likely to fail to choose a *Competitive* task when this is in fact their welfare-maximizing choice. The option to use social identity cues in belief formation enables agents with *mixed* social types to decrease the likelihood of making both types of mistakes. The multidimensionality of social identity therefore only reinforces the potential to improve decision making for the former group, not for the latter. Agents with *one-sided* social types will therefore on average fail more often to choose their welfare-maximizing task and will, consequently, have on average a lower expected utility.

The differences in the options available to form beliefs across social types affect choice behavior. An equilibrium analysis shows that differences in choice behavior across a priori identical social types can be persistent. They way intersectionality affects choice behavior at the aggregate level depends on the statistics that are available and the social constraints that agents face. The analysis of the different effects on aggregate choice behavior induced by different social constraints could be used to empirically identify the type of social constraint in a setting. An analysis of affirmative action policy shows how taking into account the intersectional effects of social context on behavior and the social constraints agents face is crucial to develop the adequate policies. I show how the design of a quota matters, and how a wrong design may simply transfer inequalities from one dimension of social identity to another. Furthermore, I shed light on the limitations of affirmative action policy, and discuss when and how informational policies could be a more effective tool to increase the representation of underrepresented groups.

Finally, the limitations of the results and the policy implications obtained in this paper are that they are derived under the assumption that agents are able to learn their optimal strategy. In reality, there may be several reasons why this may not be the case. Better knowledge about this learning process and why it may be incomplete is therefore crucial to make the step from the theoretical framework to policy in practice.

References

D. Atkin, E. Colson-Sihra, and M. Shayo. How do we choose our identity? a revealed preference approach using food consumption. *Journal of Political Economy*, 129(4): 1193–1251, 2021.

L. Beaman, E. Duflo, R. Pande, and P. Topalova. Female leadership raises aspirations and educational attainment for girls: a policy experiment in india. *Science*, 335(6068): 582586, 2012.

G. Cassan. Identity-based policies and identity manipulation: Evidence from colonial punjab. *American Economic Journal: Economic Policy*, 7(4):103–131, 2015.

G. Cassan and L. Vandewalle. Identities and public policies: Unexpected effects of political reservations for women in india. *World Development*, 143:105408, 2021.

K. Crenshaw. Mapping the margins: Intersectionality, identity politics, and violence agains women of color. *Stanford Law Review*, 43:1241–1299, 1991.

J. Denrell. Vicarious learning, undersampling of failure, and the myths of management. *Organizational Science*, 14(3):227–351, 2003.

O. Folke, L. Freidenvall, and J. Rickne. Gender quotas and ethnic minority representation: Swedish evidence from a longitudinal mixed methods study. *Politics & Gender*, 11(2):345–381, 2015.

M.M. Hughes. Intersectionality, quotas, and minority women's political representation worldwide. *American Political Science Review*, 105(3):604620, 2011.

R. Jia and T. Persson. Individual vs. social motives in identity choice: Theory and evidence from china. *ERN: Intrinsic Motivation (Topic)*, 2019.

V. Karekurve-Ramachandra and A. Lee. Do gender quotas hurt less privileged groups? evidence from india. *American Journal of Political Science*, 64(4):757–772, 2020.

C.W. Liqui Lung. On the origin and persistence of identity-driven choice behavior. Working paper, Paris School of Economics, 2022.

J. Nagel. Constructing ethnicity: Creating and recreating ethnic identity and culture. *Social Problems*, 41(1):152–176, 1994.

N. Qian and E. Nix. The fluidity of race: Passing in the united states 1880-1940. Working paper, NBER, 2015.

P. Rebughini. Agency in intersectionality. towards a method for studying the situatedness of action. *Socio*, 15:189–205, 2021.

M. Shayo. A model of social identity with an application to political economy: nation, class, and redistribution. *American Political Science Review*, 103(02):147–174, May 2009.

N. Tan. Ethnic quotas and unintended effects on womens political representation in singapore. *International Political Science Review*, 35(1):27–40, 2014.

N. Yuval-Davis. Intersectionality and feminist politics. *European Journal of Women's Studies*, 12(3):193–209, 2006.

Appendix 1: Proofs

PROPOSITION 1 (Individually Optimal Belief Formation): *Consider an agent with social type* Θ_i. *For agents with* $\alpha > \gamma$, $\sigma_i^* = \overline{\kappa}_t$ *if and only if* $\overline{\eta}_t > 1$. *Otherwise,* $\sigma^* = R$. *For agents with* $\alpha < \gamma$, $\sigma_i^* = \underline{\kappa}_t$ *if and only if* $\underline{\eta}_t < 1$. *Otherwise,* $\sigma^* = R$.

Proof. Agents choose σ_i to maximize V_i over all possible realizations of $\hat{\alpha}_i$. Consider first agents that have an $\alpha > \gamma$. The welfare-maximizing choice for these agents is $a = C$. Let γ^{σ_i} be the threshold γ_i following from playing σ_i. Now, $V_i(\sigma_i) > V_i(R)$ for $\sigma_i \neq R$ if and only if $\Phi_{\alpha,t,\sigma_i,\Pi} \geq \Phi_{\alpha,t,R,\Pi}$ for some $\sigma_i \in \{A, B\}$. Since $\Phi_{\alpha,t,\sigma,\Pi} = P(\hat{\alpha} > \gamma^{\sigma_i}|\alpha)$, this is the case when $\gamma^{\sigma_i} < \gamma^{R_i}$. This is true if and only if $\pi_x^k \geq p_x^k$ for some $k \in \{A, B\}$. When multiple social identity cues satisfy this condition, agents maximize V_i by maximizing the size of the bias, or in other words, by choosing σ_i to maximize $\gamma - \gamma^{\sigma_i}$. This is the case when they choose $\sigma_i = \overline{\kappa}_t$. When none of the social identity cues satisfy the condition, they should choose $\sigma_i = R$. Vice versa for agents with $\alpha < \gamma$, the welfare-maximizing choice is to take action $a = NC$. Therefore, $V_i(\sigma_i) > V_i(R)$ if and only if $\Phi_{\alpha,t,\sigma_i,\Pi} \leq \Phi_{\alpha,t,R,\Pi}$ for some $\sigma_i \in \{A, B\}$. This is the case if and only if $\gamma^{\sigma_i} > \gamma^R$, meaning that we need $\pi_x^k \leq p_x^k$ for some $k \in \{A, B\}$. When multiple social identity cues satisfy this condition, agents maximize V_i by maximizing the size of the bias, or in other words, by choosing σ_i to maximize $\gamma^{\sigma_i} - \gamma$. This is the case when they choose $\sigma_i = \underline{\kappa}_t$. When none of the social identity cues satisfy the condition, they should *Repress*. ■

PROPOSITION 2 (Potential to Improve Decision Making): *Simultaneous asymmetry $\pi_x^k \neq p_x^k$ along the lines of the observable characteristics θ^A and θ^B leads to inequalities in the potential to improve decision making across the different social types $\Theta = t$ with $t \in T$. Specifically, agents with mixed social types have more potential to improve decision making than agents with one-sided social types, which leads to an on average higher expected utility for agents with mixed social types.*

Proof. Agents with a *mixed* social type have both $D_t^- > 0$ and $D_t^+ > 0$. Therefore, when they have $\alpha > \gamma$, they can use social identity cues to increase the likelihood of choosing the *Competitive* task. Similarly, when they have $\alpha < \gamma$, they can use these cues to increase the likelihood of choosing the *Non-Competitive* task. Agents with a *one-sided* social type have either $D_t^- > 0$ or $D_t^+ > 0$. They can therefore either improve decision making when $\alpha > \gamma$, while they should *Repress* the use of social context in belief formation when $\alpha < \gamma$, or vice versa. Because $V(\sigma^*) > V(R)$ when $\sigma^* \neq R$, agents for whom it is optimal to not repress the use of social identity cues in belief formation have a higher expected utility. Agents with a *mixed* social type have $V(\sigma^*) > V(R)$ both when $\alpha < \gamma$ and when $\alpha > \gamma$. Agents with a *one-sided* type always have $V(\sigma^*) = V(R)$ either when $\alpha > \gamma$ or when $\alpha < \gamma$. Therefore, when we aggregate all agents with $\alpha > \gamma$ and $\alpha < \gamma$, agents with a *mixed* social type have on average a higher expected utility than agents with a *one-sided* social type. ■

PROPOSITION 3: *There exists a **Non-Neutral Regime of degree 2** in which WLOG $\pi_x^k > p_x^k$ for all $k \in \{A, B\}$. In such a regime, any possible order on the social identity cues must be such that,*

$$\pi_{\tilde{t}_{\theta'}} = \min_{k,\theta} \pi_x^k \qquad\qquad \pi_{\tilde{t}_\theta} = \max_{k,\theta} \pi_x^k$$

where $\tilde{t}_\theta$ is such that $\theta^k = \theta$, while $\tilde{t}_{\theta'}$ is such that $\theta^k = \theta'$ for all $k \in \{A, B\}$.

Proof. The social identity cues $\tilde{\pi}_1^k(\Pi, \sigma)$ induced by strategies σ and a social context Π for $k \in \{A, B\}$ are given by,

$$\tilde{\pi}_1^A(\Pi, \sigma) = \frac{S_{11} + S_{10}}{S_{11} + S_{10} + S_{00} + S_{01}} \tag{23}$$

$$\tilde{\pi}_1^B(\Pi, \sigma) = \frac{S_{11} + S_{01}}{S_{11} + S_{10} + S_{00} + S_{01}} \tag{24}$$

with $S_{x^A x^B} = p_\theta^A p_\theta^B \int \alpha \Phi_{\alpha,t,\sigma,\Pi} f(\alpha) d\alpha$ denoting the number of successful agents in the

Competitive task with social type $t = (x^A, x^B)$. We can directly infer that a *Neutral Regime* always exists. When $\pi_t = p_t$ for all $t \in T$, then, because the observable characteristics θ^A and θ^B are independently distributed, $\pi_1^A = p_1^A$ and $\pi_1^B = p_1^B$. Then, all strategies $\sigma \in \{A, B, R\}$ are equivalent. Since α and Θ are independently distributed, there will be no differences in choice behavior across social groups, and $\tilde{\pi}_1^k(\Pi, \sigma) = p_1^k$.

Now consider a perturbation of a *Neutral Regime* such that $\pi_{11}^\delta = \pi_{11} + \delta$, while $\pi_{00}^\delta = \pi_{00} - \delta$. This induces a social context such that $\pi_{11}^\delta > \pi_{00}^\delta$, while $\pi_{10}^\delta = \pi_{01}^\delta$. From Equations (22) and (23), we can infer that such a perturbation has a symmetric effect on the induced social identity cues $\tilde{\pi}_1^k(\Pi, \sigma)$ for $k \in \{A, B\}$. Let $\mathcal{S}^{sym} = \{\Pi : \pi_{11} = 1 - \pi_{00}$ and $\pi_{10} = \pi_{01}\}$ be the set of all social contexts Π in which the induced social identity cues $\tilde{\pi}_1^k(\Pi, \sigma)$ are symmetric for $k \in \{A, B\}$. Similarly, we can now define the set $\mathcal{S}_\delta^{sym} = \{\Pi : \pi_{11} = 1 - \pi_{00}, \pi_{10} = \pi_{01}$ and $\pi_{11} + \pi_{10} > \frac{1}{2}\}$, that contains all Π that are induced by a perturbation δ such that $\pi_{11}^\delta = \pi_{11} + \delta$, while $\pi_{00}^\delta = \pi_{00} - \delta$. It follows that $\mathcal{S}_\delta^{sym} \subset \mathcal{S}^{sym}$. Finally, let $\tilde{\Pi}(\sigma, \Pi_\theta)$ be a continuous function.

We now have a non-empty, compact and convex set $\mathcal{S}_\delta^{sym}$, and a continuous function $\tilde{\Pi}(\sigma, \cdot) : \mathcal{S}_\delta^{sym} \rightarrow \mathcal{S}_\delta^{sym}$.Therefore, following Brouwer's fixed point theorem, there exists a fixed point $\Pi^* \in \mathcal{S}_\delta^{sym}$ such that,

$$\Pi^* = \tilde{\Pi}(\mathcal{S}_\delta^{sym})$$

Finally, because $\Pi^* \in \mathcal{S}_\delta^{sym}$ we will have $\pi_{11}^* > \pi_{10}^* = \pi_{01}^* > \pi_{00}^*$. Therefore, $\pi_1^k > \pi_0^k$ for all $k \in \{A, B\}$ and a *Non-Neutral Regime of Degree 2* exists. That $\pi_{11}^* > \pi_{10}^* = \pi_{01}^* > \pi_{00}^*$ can be the only order that can exist in such a regime follows from Corollary 1.2. ■

COROLLARY 1.1 (One-Dimensional Lens): *For any observable characteristic θ^k such that $\pi_x^k > p_x^k$, we have population effect $\Phi_{\alpha,t:\theta^k=x,\sigma_i^*,\Pi} > \Phi_{\alpha,t:\theta^k=x',\sigma_i^*,\Pi}$ and a selection effect $E(\alpha|a = C, t : \theta^k = x) < E(\alpha|a = C, t : \theta^k = x')$. These effects are such that the order on π_x^k and $\pi_{x'}^k$ will not be reversed for any $k \in \{A, B\}$.*

Proof. Assume WLOG that $\pi_1^k > p_0^k$. Then, all agents with $\alpha > \gamma$, $\theta^k = 1$ and $k = \overline{\kappa}_t$, and all agents with $\alpha < \gamma$, $\theta^k = 0$ and $k = \underline{\kappa}_t$ will choose $\sigma = k$. Because the observable characteristics are independently distributed over the population,

$$\Phi_{\alpha>\gamma,t:\theta^k=1,\sigma,\Pi} > \Phi_{\alpha>\gamma,t:\theta^k=0,\sigma,\Pi}$$

and

$$\Phi_{\alpha<\gamma,t:\theta^k=1,\sigma,\Pi} > \Phi_{\alpha<\gamma,t:\theta^k=0,\sigma,\Pi}$$

Therefore,

$$\Phi_{\alpha,t:\theta^k=1,\sigma,\Pi} > \Phi_{\alpha,t:\theta^k=0,\sigma,\Pi}$$

Because N is arbitrarily large, this translates into population fractions.

Let $\gamma_\theta = \frac{\gamma}{\eta(\pi_x^k, p_x^k)}$. Because $\gamma_1 < \gamma$, while $\gamma_0 > \gamma$, agents with $\theta^k = 1$ choose the *Competitive* task on average for lower values of $\hat{\alpha}$ than agents with $\theta^k = 0$. Because $\hat{\alpha}$ is unbiased, this implies,

$$E(\alpha|a = C, t : \theta^k = 1) < E(\alpha|a = C, t : \theta^k = 0)$$

The number of expected successful agents is of type (α, t) is given by $P(a = C|\alpha, t)E(\alpha|a = C, t)$. Because the observable characteristics are independently distributed over the population from α, the behavior of agents in absence of asymmetry along the lines of θ^k is symmetric across the group of agents with $\theta^k = 1$ and the group of agents with $\theta^k = 0$. Let $S_{\theta^A\theta^B} = p_x^A p_x^B \int \alpha \Phi_{\alpha,t,\sigma,\Pi} f(\alpha) d\alpha$. To prove that the population and selection effects do not reverse the order on $(\pi_x^k - p_x^k)$ we need to show that, $S_1 > S_0$, where $S_1 = S_{11} + S_{10}$ and $S_0 = S_{00} + S_{01}$. This can be demonstrated by writing,

$$\begin{aligned} S_1 =& p_1^A \int_{\alpha>\gamma} \int_{\hat{\alpha}>\gamma_1} \alpha g_\alpha(\hat{\alpha}) f(\alpha) d\alpha d\hat{\alpha} + \int_{\alpha<\gamma} \int_{\gamma<\hat{\alpha}<\gamma_0} \alpha g_\alpha(\hat{\alpha}) f(\alpha) d\alpha d\hat{\alpha} + \\ & p_1^A \int_{\alpha<\gamma} \int_{\hat{\alpha}>\gamma_0} \alpha g_\alpha(\hat{\alpha}) f(\alpha) d\alpha d\hat{\alpha} \end{aligned}$$

Since,

$$S_0 = p_0^A \int_{\alpha>\gamma} \int_{\hat{\alpha}>\gamma} \alpha g_\alpha(\hat{\alpha}) f(\alpha) d\alpha + p_0^A \int_{\alpha<\gamma} \int_{\hat{\alpha}>\gamma_0} \alpha g_\alpha(\hat{\alpha}) f(\alpha) d\alpha d\hat{\alpha} + d\hat{\alpha}$$

It follows that when $p_0^A = p_1^A$ and $\gamma_1 < \gamma < \gamma_0$, then $S_1 > S_0$. ∎

COROLLARY 1.2: *Let* Π *be such that* $\pi_x^k > p_x^k$ *for all* $k \in \{A, B\}$. *Then, we have an intersectional population effect, such that* $\Phi_{\alpha,\tilde{t}_x,\sigma_i^*,\Pi} > \Phi_{\alpha,t_{mixed},\sigma_i^*,\Pi} > \Phi_{\alpha,\tilde{t}_{x'},\sigma_i^*,\Pi}$, *and an intersectional selection effect, such that* $E(\alpha|a = C, \tilde{t}_x) < E(\alpha|a = C, t_{mixed}) < E(\alpha|a = C, \tilde{t}_{x'})$. *These effects are such that the order on* π_x^k *and* $\pi_{x'}^k$ *will not be reversed for any observable characteristic* $k \in \{A, B\}$.

Proof. Consider first the agents with social type $\tilde{t}_x$. When having $\alpha > \gamma$, following Proposition 1, the σ^* of these agents always implies the strongest bias of their own noisy perception towards undertaking the task. Therefore, $\Phi_{\alpha>\gamma,\tilde{t}_x,\sigma_i^*,\Pi} = \max_{t\in T} \Phi_{\alpha>\gamma,t,\sigma_i^*,\Pi}$. When having $\alpha < \gamma$, these agents cannot use their social identity cues to bias decision making towards undertaking the outside option, and they will play $\sigma_i^* = R$. This set of agents is the only subset of all agents with $\alpha < \gamma$ that plays R, and therefore $\Phi_{\alpha<\gamma,\tilde{t}_x,\sigma_i^*,\Pi} = \max_{t\in T} \Phi_{\alpha<\gamma,t,\sigma_i^*,\Pi}$. Consequently, $\Phi_{\alpha,\tilde{t}_x,\sigma_i^*,\Pi} = \max_{t\in T} \Phi_{\alpha,t,\sigma_i^*,\Pi}$. Agents with social type $\tilde{t}_{x'}$ and $\alpha > \gamma$ belong to the only subset of agents that cannot use their social identity cues to bias decision making towards undertaking the task. Therefore, $\Phi_{\alpha>\gamma,\tilde{t}_{x'},\sigma_i^*,\Pi} = \min_{t\in T} \Phi_{\alpha>\gamma,t,\sigma_i^*,\Pi}$. When having $\alpha < \gamma$, following Proposition 1, the σ^* of these agents implies the strongest bias of their own noisy perception towards undertaking the outside option. Therefore, $\Phi_{\alpha<\gamma,\tilde{t}_{x'},\sigma_i^*,\Pi} = \min_{t\in T} \Phi_{\alpha<\gamma,t,\sigma_i^*,\Pi}$. Consequently, $\Phi_{\alpha,\tilde{t}_{x'},\sigma_i^*,\Pi} = \min_{t\in T} \Phi_{\alpha,t,\sigma_i^*,\Pi}$. Agent's with social types t_{mixed} can use their social identity cues to bias their own noisy perception both towards undertaking the task and the outside option. Following Proposition 1, $\Phi_{\alpha>\gamma,t_{mixed},\sigma_i^*,\Pi} \leq \Phi_{\alpha>\gamma,\tilde{t}_x,\sigma_i^*,\Pi}$, and $\Phi_{\alpha<\gamma,t_{mixed},\sigma_i^*,\Pi} \geq \Phi_{\alpha<\gamma,\tilde{t}_{x'},\sigma_i^*,\Pi}$. It follows therefore that $\Phi_{\alpha,\tilde{t}_x,\sigma_i^*,\Pi} > \Phi_{\alpha,t_{mixed},\sigma_i^*,\Pi} > \Phi_{\alpha,\tilde{t}_{x'},\sigma_i^*,\Pi}$. The selection effect follows from these population effects as described in the proof of Corollary 1.1, as well as the proof of the fact that the population and selection effects never reverse the order of representation in the social context. ∎

COROLLARY 2: *Consider an agents with social type* $\Theta_i = t$. *When this agents has* $\alpha > \gamma$, $\sigma_i^* = \overline{\kappa}_t$ *if and only if* $\overline{\overline{\eta}}_t > 1$. *Otherwise,* $\sigma_i^* = R$. *When agents have* $\alpha < \gamma$, *then* $\sigma_i^* = \underline{\kappa}_t$ *if and only if* $\underline{\underline{\eta}}_t < 1$. *Otherwise,* $\sigma_i^* = R$.

Proof. When agents choose σ_i to maximize $V(\sigma_i)$, it follows that they will choose $\sigma_i = F$ if and only if $F = \text{argmax}_{\sigma_i\in\{A,B,F,R\}} V(\sigma_i)$. When $\alpha > \gamma$, maximizing $V(\sigma_i)$ is equivalent to maximizing $\Phi_{\alpha,t,\sigma_i,\Pi}$. We only have $\Phi_{\alpha,t,F,\Pi} = max_{\sigma_i\in\{A,B,F,R\}}\Phi_{\alpha,t,\sigma_i,\Pi}$ when $\pi_t > p_t$ and $F = \overline{\kappa}_t$. Vice versa, when $\alpha < \gamma$. ∎

COROLLARY 3: *The introduction of the strategy $\sigma_i = F$ does not invalidate the existence of a Non-Neutral Regime of Degree 2.*

Proof. Consider a social context $\Pi \in \mathcal{S}^{sym}$ and perturb this social context, such that $\Pi^\delta \in \mathcal{S}^{sym}_\delta$. If $F = \text{argmax}_{t,k\in\{A,B\}}\, \eta(\pi_t, p_t), \eta(\pi^k_x, p^k_x)$ for some $t \in \{10, 01\}$, then $\tilde{\pi}_{01} \neq \tilde{\pi}_{10}$ and $\tilde{\Pi}(\sigma, \Pi^\delta) \notin \mathcal{S}^{sym}_\delta$. Therefore, $\tilde{\Pi}(\sigma, \Pi^\delta) \in \mathcal{S}^{sym}_\delta$ if and only if $k = \text{argmax}_{t,k\in\{A,B\}}\, \eta(\pi_t, p_t), \eta(\pi^k_x, p^k_x)$ for all $t \in \{10, 01\}$. When this condition is met, we can again use Brouwer's fixed point theorem to show that there exists a fixed point Π^* such that $\Pi^* = \tilde{\Pi}(\sigma, \Pi^*)$. ■