UNIVERSITY OF CAMBRIDGE
Faculty of Economics

JANEWAY INSTITUTE CAMBRIDGE

CAMBRIDGE WORKING PAPERS IN ECONOMICS
JANEWAY INSTITUTE WORKING PAPERS

On the Origin and Persistence of Identity-Driven Choice Behavior

Caroline W. Liqui-Lung
University of Cambridge

Abstract

Recent literature shows equally performing individuals belonging to different social groups can have different beliefs about their abilities, and, consequently, make different educational and occupational choices. This paper contributes to understanding this phenomenon. I show people can use statistics about the prevalence of their social group among the successful individuals in a task to cope with the adverse effects of momentary sources of noise and improve decision making on average, even when these statistics are irrelevant in a Bayesian sense. This individually optimal behavior can nevertheless induce persistent asymmetries in belief formation and choice behavior across otherwise identical social groups.

Reference Details

2274	Cambridge Working Papers in Economics
2232	Janeway Institute Working Paper Series
Published	15 December 2022
Revised	16 August 2023
JEL codes	D81, D91, Z13
Websites	www.econ.cam.ac.uk/cwpe
	www.janeway.econ.cam.ac.uk/working-papers

On the Origin and Persistence of Identity-Driven Choice Behavior

Caroline W. Liqui-Lung *

Version: August 16, 2023

Recent literature shows equally performing individuals belonging to different social groups can have different beliefs about their abilities, and, consequently, make different educational and occupational choices. This paper contributes to understanding this phenomenon. I show people can use statistics about the prevalence of their social group among the successful individuals in a task to cope with the adverse effects of momentary sources of noise and improve decision making on average, even when these statistics are irrelevant in a Bayesian sense. This individually optimal behavior can nevertheless induce persistent asymmetries in belief formation and choice behavior across otherwise identical social groups.

(JEL: D-81, D-91, Z-13)

*University of Cambridge (Email: cll64@cam.ac.uk). I would like to thank Olivier Compte, Philippe Jehiel, Roland Bénabou, Debraj Ray, Alessandra Casella, Abhijit Banerjee, Muhamet Yildiz, Vincent Crawford, Willemien Kets, Dilip Mookherjee, Gabrielle Fack, Jean-Marc Tallon and the participants of ESSET, SITE, ERINN, ThReD, the Royal Economic Society Annual Conference, the North American, European, Asian and China Meetings of the Econometric Society, ADRES, the ECINEQ annual conference, the Warwick Economics PhD Conference and the seminars at MIT, PSE, the University of Cambridge, INSEAD and the University of Amsterdam (CREED) for their useful comments and suggestions.

1 Introduction

Equally performing individuals belonging to different social groups can have systematically different beliefs about their own abilities. Recent studies show this for male and female math students in Germany (Lippmann and Senik, 2018), black and white college students in the US (Hodge et al., 2008), men and women in the labor market (Exley and Kessler, 2022), students from different castes in India (Mukherjee, 2017), and socially more and less advantaged high school children in France (Guyon and Huillery, 2021). The last paper identifies this differential belief formation moreover as one of the main drivers of why socially less advantaged high school children are less likely to pursue elite educational pathways than their equally performing, but socially more advantaged peers. Such identity-driven choice behavior contributes to inequalities across social groups and negatively affects social diversity (Blau and Kahn, 2017). To develop the adequate policy to promote equality, social mobility and diversity across social groups, we therefore need to understand how and why such differential belief formation arises.

This paper contributes to that objective by showing why people may find it optimal to let their beliefs about their ability in a task be influenced by informationally irrelevant statistics about which social groups are relatively more successful, and how this can lead to persistent differences in belief formation and choice behavior across a priori identical subgroups. The story is as follows. Although we may generally have an accurate perception of our abilities, exogenous factors, such as emotions or recent feedback, may make us momentarily too optimistic or pessimistic (see e.g. Fiedler and Bless (2000) and Elster (1996)). This induces noise in our perception that makes us prone to making mistakes when choosing whether to undertake tasks related to these abilities.

The effect of this noise on decision making is asymmetric. To illustrate, consider high school students that decide on entering a math competition. For a student who is not strong in math, having a bad day and being momentarily too pessimistic will not affect choice behavior, since she will refrain from entering the competition either way. On the other hand, if the student is momentarily too optimistic, this makes her prone to making the type I error of entering the competition, while this is not optimal. For a student who is very good at math, the opposite is true. When this student is too pessimistic, she is prone to making the type II error of not entering the competition, while this would have been optimal.

Now, assume students observe data about a pool of students from the previous year and that male students were relatively overrepresented among those successful in the competition. I pose a male student can use this information to make himself believe that, because relatively more male students were successful, he is more likely to be successful himself, while a female student can make herself believe that, because women are relatively underrepresented among those successful, she is less likely to be successful.[1] Consequently, I show how students can use the belief that these statistics are relevant for their own chances of success to cope with the earlier described adverse effects of momentary noise on decision making. To illustrate, take a high-ability male student. His rational self may know he is generally strong in math, and he may be aware he is prone to making a type II error induced by negative emotions in the moment. To decrease the likelihood he will

[1] This is in line with research in social psychology, such as the work of Seligman (2006) who shows people can take the successes and failures of others like them as evidence they will fail or succeed as well, and Murden (2020) who shows how our behavior is influenced by the choices and outcomes of others.

make such an error, he can boost his own noisy perception upwards by believing the outcomes of other male students are relevant to his own chances of success.[2] A high-ability female student, on the other hand, would only increase the chances of making a type II error when she believes the underrepresentation of women is relevant. She may have learned from experience with other math-related tasks that she refrains too much from undertaking these tasks when she believes her gender is relevant, and will therefore ignore the statistic. This is in line with Pronin et al. (2004), that shows how strong female math students actively disidentify with character traits that are believed to be strongly related with the negative math-gender stereotype. Similarly, low-ability female students can decrease the likelihood of making a type I error by biasing their noisy perception downward with the statistic, while low-ability male students should ignore the data.

As a result, male students will on average be more optimistic about their chances of success in the competition, while female students will on average be more pessimistic. This illustrates how social context can cause differences in belief formation across a priori identical individuals, even when it is irrelevant in a Bayesian sense. This is in line with results of experiments showing that, on average, men are more overconfident than women in fields with a strong male connotation, while the opposite is true in fields with a strong female connotation (e.g. Coffman (2014) and Flory et al. (2015)). Furthermore, male students will be more likely to make type I errors, while female students will be more likely to make type II errors. Hence, more male than female students will enter the math competition. This reinforces the overrepresentation of male students among those successful. The story

[2]This is in line with the more popular concepts of ‘life hacks’ or ‘coping strategies’(Peters, 2021)

can therefore explain the differential belief formation in the research cited earlier, and shows how it can induce persistent identity-driven choice behavior, even when there are no ability differences across the different groups. Furthermore, different types of statistics induce different aggregate choice behavior, and we could avoid the persistence of identity-driven choice behavior by influencing the data available and the structure that is put on this data. These insights provide implications for informational policies to achieve social diversity and fight harmful stereotypes.

I propose a model in which both social context and its effect on belief formation are determined endogenously, without assuming social context directly affects utility. Agents choose between an *ability-driven* task with an individual-specific probability of success, of which they only have a noisy, but unbiased perception, and an outside option with a known probability of success that is the same for all agents.[3] An agent's type also includes an observable characteristic that determines their social group. They observe *social identity cues* that stem from the prevalence of their subgroup among the already successful individuals in the task. To ensure social context is informationally irrelevant, the characteristic and the individual-specific probability of success are independently distributed over the population. In the spirit of Compte and Postlewaite (2019), agents choose between two subjective belief-formation processes; one in which they naively follow their noisy perception, and one in which this noisy perception is influenced by their *social identity cue* in a direction contingent on their social type. Agents choose between the task and the outside option to maximize subjective expected utility.

[3]This noise should be interpreted as the effects of momentary emotions or distractions. If agents were able to switch this noise off, they would behave as Bayesians. To show that a systematic bias is not what drives the results in this model, agents are not systematically over- or under confident. Also, we obtain similar results when the perception of chances of success in both the *ability-driven* task and the outside option are noisy.

Hence, choices of belief-formation rules induce choices of tasks that in turn give rise to social identity cues. To study the mutually stable choices of belief formation and tasks in this process, I use a static solution concept. I define a fitness criterion that states a belief-formation rule is optimal when it maximizes expected utility on average over all possible realizations of the agent's noisy perception.[4] I then analyze the fixed points in the social identity cues induced by these individually optimal strategies.[5]

The idea that equilibrium is the result of people attempting to choose among strategies according to a fitness value is standard. The non-standard aspects of this model are that the set of strategies represents a limited set of belief-formation processes, and that equilibrium beliefs are disciplined in a manner that is different from concepts such as the Berk-Nash Equilibrium (Esponda and Pouzo, 2016), the Self-Confirming Equilibrium (Fudenberg and Levine, 1993) or the Personal Equilibrium (Spiegler, 2016). In the latter concepts, equilibrium beliefs are consistent with observational feedback, ensuring they are closest to the truth. In this paper, beliefs are consistent with fitness, allowing agents to make decisions that are better aligned with welfare maximization. As in Compte and Postlewaite (2004) and Brunnermeier and Parker (2005), agents deviate from traditional Bayesian beliefs when this enhances expected utility.

[4]One foundation for this behavior is in the spirit of Benabou and Tirole (2006). A calm, rational self knows the true underlying type, while an impulsive self can be biased in the moment. The rational self ties the hands of the 'in the moment' self to limit its adverse effects on choice. Another foundation can be that the agent compares the two belief formation heuristics, and learns from experience with similar tasks throughout life, through reinforcement learning or sampling, which heuristic leads to on average more successful outcomes, without having to fully understand the relationship between his choice of belief formation rule, choice of task and outcome.

[5]Dekel et al. (2007) use a similar approach, but in the latter, preferences are determined in a dynamic process according to their fitness with respect to the preferences in the rest of society, while in this paper, preferences are determined according to their fitness with respect to a certain task.

I show how informationally irrelevant social cues become valuable when their use generates a bias towards the welfare-maximizing choice of task. Current economic literature models the effect of social identity on decision making predominantly through the introduction of direct identity-related utility derived from for example self-image or the fear of being punished by peers (Akerlof and Kranton, 2000), from representation in society (Carvalho and Pradelski, 2022), through costly interaction with people different from yourself (Battu et al., 2007) or through status and perceived similarity (Shayo, 2009). Another approach can be found in the literature on discrimination (Onuchic, 2023), affirmative action (Benhabib et al., 2010) and social pressure (Bursztyn and Jensen, 2017), where the effect of social identity on decision making is driven by strategic interaction between agents. Finally, the literature on social learning shows how agents can learn form the choices or outcomes of others when this information is relevant in a Bayesian sense (e.g. Banerjee (1992) and Wolitzky (2018)). This model provides a different, but possibly complementary view, where the use of social cues allows agents to manage the degree of over- or under-confidence regarding their chances of success through a distinct processing of their noisy perception. The optimal use of social cues results in the optimal management of confidence to improve decision making on average. This approach could be extended to situations in which there is a real value to biased confidence, as in Compte and Postlewaite (2004), Brunnermeier and Parker (2005) or Benabou and Tirole (2002).

I show the existence of a stable population equilibrium in which task allocation and belief formation differ between a priori identical subgroups. This is particularly enabled by a relatively attractive outside option. We would therefore expect social cues to especially drive choice behavior for tasks where few people try and succeed, like executive positions, sports or top educational programs. Furthermore, differences in choice behavior are no longer persistent when people process within-group success rates.

These results shed light on how stereotypes (Bordalo et al., 2016) or social norms (Akerlof and Kranton, 2000) can arise endogenously, and how this can be driven by the particular statistics people take into account and the data they have access to. Furthermore, Hoff and Stiglitz (2010) assume that, when groups have been historically treated as inferior, this affects how they interpret failure, which affects utility directly through performance. This paper shows how such belief formation may arise endogenously, without it affecting performance. Also, the model shows how differences in beliefs across a priori identical groups, like in Piketty (1995), Benabou and Tirole (2006), Frick et al. (2018) and Peski and Szentes (2013) can be collectively sustained and constitute an equilibrium without introducing direct interaction between agents, nor having a common state of the world or assuming direct effects of beliefs of others on an agent's preferences. Finally, the model shows how agents may want to adhere to a specific self-image (Akerlof and Kranton, 2000) or mental model (Hoff and Stiglitz, 2016) in a particular social context because of its instrumental value in decision making. In Liqui-Lung (2023), I particularly focus on this question. I extend the model with multi-dimensional social identities and discuss why agents may want to focus on the statistic related to one particular dimension of their social identity compared to another. This provides insights regarding why different social identities may be salient to the same people in different social contexts.

Furthermore, the differential use of social identity cues in belief formation induces both a difference in the propensity to choose the *ability-driven* task across subgroups, and a difference in mean competence: agents belonging to the socially less successful subgroup have a lower propensity to choose the *ability-driven* task, but, conditional on choosing this task, they tend to be more competent on average than agents belonging to the socially more successful group.[6] These effects are also obtained in several models of statistical discrimination and affirmative action (in the style of e.g. Coate and Loury (1993) and Phelps (1972)), and in that light, this model could be interpreted as a model of optimal self-discrimination, where there is no interaction between agents[7]. Finally, the influence of social context on beliefs especially drives choice behavior of agents with average ability levels. This could explain why Buser et al. (2014) find that the gender gap in curriculum choice shows up precisely at the mean: while average men choose highly mathematical curricula, average women choose very humanities-intensive curricula, which causes women to be over-represented in the latter, and men overrepresented in the former.

The rest of the paper is organized as follows. Section 2 introduces the model. Section 3 presents the results. Section 4 discusses informational policy implications and Section 5 discusses the main assumptions of the model. Section 6 reviews the general literature on the topic, and Section 7 concludes. All formal proofs can be found in the appendix.

[6] A study by S&P market intelligence shows that men outnumber women in the CFO job by about 6.5 to 1. Companies appointing female CFO's saw nevertheless a 6% increase in profits and an 8% better stock return compared to companies appointing male CFO's. Moreover, female CFO's brought in $1.8 trillion of additional cumulative profit and therefore significantly outperformed their male peers. This result is also in line with Niederle and Vesterlund (2007), who show that too few high-skilled women and too many low-skilled men enter competitive math-related tasks.

[7] I discuss in the paper how the presence of discrimination would reinforce the results.

2 The Model

2.1 The Environment

I consider a society with $i = 1, ..., N$ agents, with N arbitrarily large. Each agent chooses an action $a \in \{C, NC\}$, where C and NC represent classes of tasks of respectively a *Competitive* and a *Non-Competitive* type. The outcome of a 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.

Noisy Perceptions - The probability α_i is unobservable, and agents only have a noisy perception $\hat{\alpha}_i$ regarding their own probability of success.[8] To show a systematic bias in belief formation is not the mechanism that drives the results in this model, I assume this noisy perception is unbiased. Consequently, I pose $\hat{\alpha}_i$ stems from a distribution g_{α_i} with $E(\hat{\alpha}_i) = \alpha_i$.

Social Context - Agents have an *observable characteristic* that represents for example their gender, ethnicity or social class. This characteristic is public information. To simplify the exposition of the model, I denote this characteristic by a binary variable θ_i with realizations $x \in \{0, 1\}$. Hence, each agent is fully described by her type $\{\alpha_i, \theta_i\}$. I let p_x be the fraction of the population with an observable characteristic $\theta_i = x$. To isolate the mechanism through which social identity affects choice behavior in this model, I assume the probability α and the observable characteristic θ are independently distributed over the population.[9] Agents have access to public data that consists of the outcome variables and observable characteristics of agents that have already made the choice. Society typically structures this information. To illustrate what drives the results in this model, in this section I focus on one particular statistic. In the Section 4.1, I discuss how different data and different structures on information affect behavior.

Let $\mathcal{N}_{C,x} = \{i \in N, \theta_i = x, a_i = C\}$ be the set of all individuals of type $\theta_i = x$ that have chosen the *Competitive* task. Let $\mathcal{N}_C = \{i \in N, a_i = C\}$ be the set of all individuals that have chosen the *Competitive* task, which implies $\mathcal{N}_{C,x} \subset \mathcal{N}_C$.

[8]As discussed in the introduction, this noise should be interpreted as the effects of momentary emotions or distractions. If agents were able to switch this noise off, they would behave as Bayesians.

[9]In section 5.3, I discuss how discrimination and other direct effects of social identity on utility interact with the mechanism presented in this paper.

I pose that society provides the statistic,

$$\pi_x = \frac{\sum_{i \in \mathcal{N}_{C,x}} Y_i}{\sum_{i \in \mathcal{N}_C} Y_i}$$

which is the fraction of successful individuals with characteristic $\theta = x$ among all successful individuals in the *Competitive* task. I call this fraction π_x the '*social identity cue*' for an agent with observable characteristic $\theta_i = x$. The '*social context*' of the population is defined as the vector $\Pi = (\pi_x)_{x \in \{0,1\}}$. Because α and θ are independently distributed over the population, this '*social context*' is not relevant to agents in a Bayesian sense. Instead, I introduce the option to agents to use π_x to bias their noisy perception $\hat{\alpha}_i$.

Subjective Belief Formation - I model agents that have an imperfect idea about their economic environment and think that using '*social context*' could be useful to form a belief about their probability of success α_i, even if they are not a priori sure of that. Specifically, I assume agents have a natural 'urge' to look at others like them, and they have the option to either *Repress* or *Not Repress* this urge. I introduce the following family of belief formation processes with which agents form a subjective belief $\hat{p}_i$ about their probability of success α_i, and assume agents have some discretion in finding out which belief formation process suits them best. For any value $\pi, p \in [0, 1]$, let η be a 'response function' that is non-decreasing, such that

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

Furthermore, let $\eta_x = \eta(\pi_x, p_x)$. Agents choose a strategy $\sigma_i \in \{R, NR\}$, and

$$\hat{p}_i = \begin{cases} \hat{\alpha}_i & \text{if } \sigma_i = R \\ \eta_x \hat{\alpha}_i & \text{if } \sigma_i = NR \end{cases} \tag{4}$$

Depending on whether agents let their belief formation be influenced their social cue[10], their subjective belief can take two values; $\hat{p}^R$ or $\hat{p}^{NR}$. With a subjective Bayesian interpretation in mind, $\sigma_i = R$ corresponds to a world view in which private and observable characteristics are uncorrelated, while $\sigma_i = NR$, corresponds to a view in which the two are correlated, with (π_x, p_x) informing about the sign and strength of that correlation[11]. When $\sigma_i = NR$, the agent biases her noisy perception $\hat{\alpha}_i$ in a direction contingent on her social type. If the agent belongs to the socially more successful subgroup, this belief-formation process leads to an optimistic interpretation of $\hat{\alpha}$, while this leads to a pessimistic interpretation when the agent's subgroup is underrepresented among the successful individuals.

[10]I do not model precisely the agent's thought processes leading to these two 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 can be conductive to the phenomenon I mean to describe.

[11]In this case, agents make themselves believe that the probability of success of agents like them has some predictive value for their own probability of success, and the model can be interpreted as agents exhibiting an attribution error.

Subjective Utility Maximization - Agents derive utility from being successful and the utility function can be represented by $u_i = Y_i$. Each agent chooses her action a_i to maximize $E(u_i)$ given her subjective belief $\hat{p}_i^\sigma$, and will choose the *Competitive* task if and only if $\hat{p}_i^\sigma > \gamma$. One could say therefore that agents are subjectively 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 agent's subjective belief $\hat{p}_i^\sigma$, where $\hat{p}_i^\sigma \in \{\hat{p}_i^R, \hat{p}_i^{NR}\}$. 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 two thresholds, such that agent i chooses $a = C$ if and only if $\hat{\alpha}_i > \gamma_i$, where

$$\gamma_i = \begin{cases} \gamma & \text{when } \sigma_i = R \\ \frac{\gamma}{\eta_x} & \text{when } \sigma_i = NR \end{cases} \tag{5}$$

The strategy '*Not Repress*' implies therefore 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 also be directly specified as the choice set $\gamma_i \in \{\gamma, \frac{\gamma}{\eta_x}\}$. This choice set can be different for agents with different values of θ, which will be the key driver of the equilibrium results.

2.2 The Solution Concept

Choices of belief formation affect choices of tasks. This leads to outcomes that induce cues that in turn affect belief formation. To tractably capture the fixed points in this dynamic process, I use a static solution concept in which I assume that, given social context, agents choose a strategy σ according to a fitness value that I define below. A population equilibrium is then a fixed point in social context that is induced by the optimal strategies. 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 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 social context arise from agents belonging to different generations making a specific choice of task once in their lifetime.

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

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

This probability Φ follows from the distribution $g_{\alpha_i}(\hat{\alpha}_i)$ given the choice of strategy σ_i. From an outsiders perspective, the expected pay-off for agent i of type $\{\alpha, x\}$ playing σ_i given Π over all possible realizations of $\hat{\alpha}$ is,

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

with $\sigma_i \in \{R, NR\}$. I then define individual optimality as follows.

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

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

Individual optimality implies that an agent chooses her belief formation to maximize her expected pay-off on average over all possible realizations of $\hat{\alpha}_i$. The fitness value of a strategy σ is determined by both the agent's type $\{\alpha, x\}$ and the social context Π.

I assume agents compare $V_i(R)$ and $V_i(NR)$, and choose their strategy σ_i according to Definition 1. This assumption can be justified with the view that agents learn their optimal belief formation from their own experience with similar choices through for example reinforcement learning or a sampling process[12]. The true probability α_i determines the outcomes agents observe, which enables them to learn whether it is optimal to *Repress* without precise knowledge of the relationship between the choice of strategy, choice of task and the observed outcome. Because the set of strategies is small, this is easy for agents to calculate.[13] An alternative foundation can be in the spirit of Benabou and Tirole (2006). A calm, rational self knows the true underlying type, while an impulsive self can be biased in the moment. The first self may decide on a belief formation rule, while the

[12]The dynamic story underlying the reduced-form analysis is that agents make similar ability-related choices throughout their lifetime. For example, early in life they choose whether to 'undertake a math-related major', while later in life they choose whether to 'pursue a STEM career'.

[13]It seems plausible that if agents are able to learn their optimal strategy σ conditional on α, they should also be able to retrieve their true value of α from this optimal strategy. This line of thought is nevertheless driven by the simplification of the model in which α and γ are fixed over the lifetime of an agent, and I will elaborate more in the section 5.2 on how the model can account for sophisticated agents that understand what their fitness signals about their true probability of success.

second self chooses an action at a given point in time, given the noisy perception and the earlier chosen belief formation rule.

Finally, one could say that agents are boundedly rational in the sense that not all belief formation processes can be compared. This aspect of bounded rationality should be considered a modelling device that helps to keep the model parsimonious. Because of the simplifying assumption that α and γ are fixed, and because α and θ are independently distributed, the analysis would be degenerate if agents could compare all possible functions of $\hat{\alpha}$ and π_x. The key insight from the model is that it shows the difference with a Bayesian model, by analyzing whether, when agents do not have the tools to correct for this type of noise, this can open the door for them to use information that is irrelevant, but that could still improve decision making. We shall see in Section 5 how the results extend to the more realistic case in which α and γ vary or in which such learning would be imperfect.

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

$$\tilde{\pi}_x(\sigma, \Pi) = \frac{p_x \int \alpha \Phi_{\alpha,x,\sigma,\Pi} f(\alpha) d\alpha}{\sum_{x \in \{0,1\}} p_x \int \alpha \Phi_{\alpha,x,\sigma,\Pi} f(\alpha) d\alpha} \tag{8}$$

where $f(\alpha)$ is the probability density function of α and $\tilde{\pi}_x(\sigma, \Pi)$ 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,*

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

In other words, a population equilibrium is a fixed point in '*social context*' when all agents play their individually optimal strategy.

3 The Results

3.1 On the Origin of Identity-Driven Choice Behavior

Example - Consider a firm in which agents choose whether to pursue a career in management (C) or a clerical job (NC). Assume these agents observe the current pool of successful managers, and that women are relatively overrepresented in this pool. Let $\theta_i = 1$ denote being a woman and assume $p_1 = p_0$. Let $\hat{p}_i^\sigma(\alpha, x)$ be the subjective belief $\hat{p}_i^\sigma$ implied by an agent of type $\{\alpha, x\}$ playing strategy σ_i. To illustrate behavior, consider agents of type $\alpha > \gamma$. The welfare-maximizing choice for these agents is to pursue a management career. To maximize expected utility, their belief formation should be chosen to maximize the likelihood they choose this career over all possible realizations of $\hat{\alpha}$. Hence, it is optimal to *Repress* the urge to look at others when $P(\hat{p}_i^{NR}(\alpha, x) > \gamma) \leq P(\hat{p}_i^R(\alpha, x) > \gamma)$. Figure (1) shows the different probabilities $P(\hat{p}_i^\sigma(\alpha, x) > \gamma)$ for $\sigma \in \{R, NR\}$ and $x \in \{0, 1\}$.

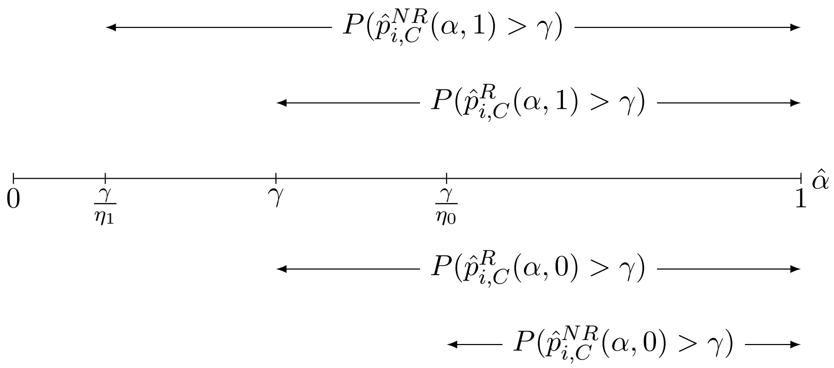

Figure 1: To probabilities to choose $a = C$ for an agent with $\alpha > \gamma$ in a social context such that $\pi_0 < \pi_1$.

Because women are overrepresented among the currently successful managers in the firm, *Not Repressing* the use of the social identity cue in belief formation causes female agents to deflate the threshold above which they think they are 'good enough' to become a successful manager. Consequently, choosing a management career becomes relatively more attractive than choosing a clerical job, and hence female agents with $\alpha > \gamma$ should choose *Not Repress* to improve decision making on average. For male agents with $\alpha > \gamma$, the opposite holds. Because men are underrepresented among the successful managers, the strategy *Not Repress* would inflate the threshold above which they think they are 'good enough'. This would make a management career relatively less attractive. Male agents with $\alpha > \gamma$ should *Repress* the urge to look at the outcomes of other men. The opposite reasoning applies to agents with $\alpha < \gamma$. In general, Proposition 1 shows that agents with $\alpha > \gamma$ will want to take their social identity into account when they belong to a socially more successful group, while they will wish to avoid it when they belong to a socially less successful group, and vice versa for agents with $\alpha < \gamma$. This shows how choice behavior can be driven by social context, even when it is irrelevant in a Bayesian sense and has no direct effect on utility.

PROPOSITION 1 (Individually Optimal Belief Formation): *The individually optimal strategies σ^* given an agent's type $\{\alpha, x\}$ are the following:*

- *The individually optimal strategy σ^* is 'Not Repress' for agents of type $\{\alpha, x\}$ such that $\alpha > \gamma$ and $\pi_x > p_x$ or $\alpha < \gamma$ and $\pi_x < p_x$*
- *The individually optimal strategy σ^* is 'Repress' for agents of type $\{\alpha, x\}$ such that $\alpha > \gamma$ and $\pi_x < p_x$ or $\alpha < \gamma$ and $\pi_x > p_x$*

Talent will always find its way - The ability to improve decision making on average using social identity cues is a function of the true probability α. Specifically, if we assume the variance of $\hat{\alpha}$ is uncorrelated with α, the use of social identity cues in belief formation is on average most beneficial to those who have a true probability of success α close to γ, while agents with extremely low or extremely high values of α are always more likely to make the correct choice, independent of their observable characteristics and the social context in which they make their decisions. In this model, social context therefore especially affects choice behavior of agents with α close to γ, while it has little effect on agents with extreme values of α.

3.2 On the Persistence of Identity-Driven Choice Behavior

When one subgroup is overrepresented among the successful individuals, this affects both how many and what type of agents choose the *Competitive* task. Specifically, when $\theta_i = x$ implies a more pessimistic processing of the noisy perception $\hat{\alpha}_i$, then those who choose the *Competitive* task despite this, tend to have a larger success rate on average than those who choose this task with a characteristic implying an optimistic processing of $\hat{\alpha}_i$. This is what we call the '*selection effect*'. On the other hand, the population of those that belong to the socially less successful subgroup choosing the *Competitive* task tends to be smaller than the population of those choosing the task belonging to the socially more successful subgroup. This is what we call the '*population effect*'. Corollary 1 formalizes this.

COROLLARY 1: *Let $x' \in \{0,1\}$ be the complement of x and assume WLOG that $\pi_x > \pi_{x'}$. The optimal use of social identity cues has both a population effect, such that $\Phi_{\alpha,x,\sigma_i,\Pi} > \Phi_{\alpha,x',\sigma_i,\Pi}$ and a selection effect, such that $E(\alpha|a = C, x) < E(\alpha|a = C, x')$. The strength of both effects is such that the order $\pi_x > \pi_{x'}$ will always be preserved.*

Example - Proposition 1 implies women with $\alpha > \gamma$ choose *Not Repress*, while men with $\alpha > \gamma$ choose *Repress*. Similarly, men with $\alpha < \gamma$ choose *Not Repress*, while women with $\alpha < \gamma$ choose *Repress*. Figure (2) shows the probabilities $P(\hat{p}_i^{\sigma}(\alpha, x) > \gamma)$ for $\sigma = \sigma^*$ and $x \in \{0,1\}$.

Since $\frac{\gamma}{\eta_1} < \gamma$ and $\gamma < \frac{\gamma}{\eta_0}$, women are more likely to, both optimally and suboptimally, pursue a management career. This demonstrates the '*population effect*'. Because the noisy perception is unbiased, higher realizations of $\hat{\alpha}$ are more likely for agents with higher true probabilities α. Consequently, women choose the management task for on average lower realizations of $\hat{\alpha}$ than men and $E(\alpha|a = C, 1) < E(\alpha|a = C, 0)$. This demonstrates the '*selection effect*'. The selection and population effect will not reverse the order on (π_1, π_0). Because α

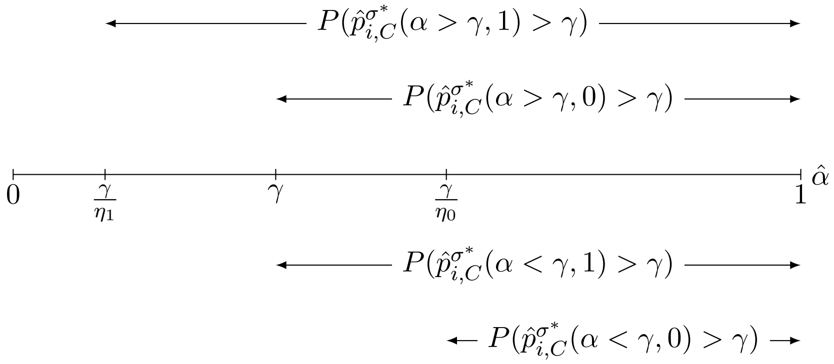

Figure 2: The probabilities $P(\hat{p}^{\sigma}_{i}(\alpha, x) > \gamma)$ for $\sigma = \sigma^*$ and $x \in \{0, 1\}$

and θ are independently distributed, the fraction of women with $\hat{\alpha} > \gamma$ in an arbitrary large population is equal to the fraction of men with $\hat{\alpha} > \gamma$. The population of women with $\alpha > \gamma$ choosing $a = C$ consists therefore of women with $\hat{\alpha} > \gamma$ plus women with $\frac{\gamma}{\eta_1} < \hat{\alpha} < \gamma$, while the population of men with $\alpha > \gamma$ choosing $a = C$ only consists of men with $\hat{\alpha} > \gamma$. Therefore, even though men have on average a higher success rate conditional on choosing $a = C$, the expected number of successful women will be larger than the expected number of successful men.

Existence and Stability of Population Equilibria - Whether identity-driven choice behavior can be persistent, depends on whether the population and selection effect reinforce or shrink differences between π_1 and π_0. Definition 3 defines the two scenarios that could appear in equilibrium.

DEFINITION 3 (Equilibrium Regimes): *In a '**Neutral Regime**' the allocation of individuals over tasks is symmetric across different subgroups, and $\pi_x = p_x$. In a '**Non-Neutral Regime**' the allocation of individuals over tasks is asymmetric across different subgroups, and $\pi_x \neq p_x$.*

Example - Consider again the example of a firm where male and female agents choose between pursuing a management career and a clerical job. Now, also assume agents have the following extreme response function,

$$\eta(\pi, p) = \begin{cases} +\infty & \text{if } \pi > p \\ 1 & \text{if } \pi = p \\ -\infty & \text{if } \pi < p \end{cases} \tag{10}$$

When $\pi_0 = \pi_1 = \frac{1}{2}$, the strategies *Repress* and *Not Repress* are equivalent, and this social context induces social identity cues such that $\tilde{\pi}_1(\sigma^*, \frac{1}{2}) = \frac{1}{2}$. In other words, a '*Neutral Regime*' always exists. Nevertheless, as soon as agents observe slightly more women than men among the successful managers, such that $\pi_1 > \pi_0$, the extreme response function $\eta(\pi, p)$ causes all women with $\alpha > \gamma$ to choose to pursue a management career, while all men with $\alpha < \gamma$ will choose the clerical job. Consequently, $\tilde{\pi}_1(\sigma^*, \pi_1) > \pi_1$, while $\tilde{\pi}_0(\sigma^*, \pi_0) < \pi_0$ and the 'Neutral Regime' becomes unstable.

Using this extreme case, we can show that any induced social identity cue $\tilde{\pi}_1(\sigma, \Pi)$ is always bounded from above. Specifically,

$$\frac{\tilde{\pi}_1(\sigma, \Pi)}{\tilde{\pi}_0(\sigma, \Pi)} \leq \frac{p_1 \int_{\alpha<\gamma} \int_{\hat{\alpha}>\gamma} \alpha g_\alpha(\hat{\alpha}) f(\alpha) d\hat{\alpha} d\alpha + p_1 \int_{\alpha>\gamma} \alpha f(\alpha) d\alpha}{p_0 \int_{\alpha>\gamma} \int_{\hat{\alpha}>\gamma} \alpha g_\alpha(\hat{\alpha}) f(\alpha) d\hat{\alpha} d\alpha} \tag{11}$$

Proposition 2 shows a sufficient condition for the existence of a stable '*Non-Neutral Regime*' obtained through analyzing when a '*Neutral Regime*' becomes unstable.

PROPOSITION 2 (Existence Non-Neutral Regime): *Let* $p_0 = p_1$*, and let* $\delta > 0$ *be a small value with which we disturb a 'Neutral Regime'. A sufficient condition for the co-existence of a stable 'Non Neutral Regime' with a 'Neutral Regime' is as follows,*

$$\frac{\partial \gamma^{NR}}{\partial \delta} \frac{|\frac{\partial S}{\partial \gamma}|}{S} > 4 \tag{12}$$

where $\gamma^{NR} = \frac{\gamma}{\eta_x}$ *and* $S = \int \alpha G_\alpha(\gamma) f(\alpha) d\alpha$*.*

Proposition 2 shows the two ingredients that contribute to making a '*Neutral Regime*' unstable. First of all, a change of δ in Π away from a 'Neutral Regime' must have a sufficiently large effect on the choice behavior of agents at the individual level. Specifically, the induced change in the threshold γ^{NR} of agents that choose to '*Not Repress*' must be large enough. This change depends, first, on the derivative of the response function $\eta(\pi_x, p_x)$ at the '*Neutral regime*'. Secondly, because of the linearity of γ^{NR} in γ, this change is multiplicative in γ. The other ingredient that contributes to the instability of a '*Neutral Regime*' is driven by the effect of a change δ on the outcomes at the aggregate level. This is captured by the elasticity of S in γ, where S is the total number of successful people at the *Competitive* task. The absolute value of this elasticity is increasing in γ, since the more attractive the outside option, the lower the number of agents S that tries the *Competitive* task. Moreover, the higher γ, the higher the success rate of agents that choose this task. Consequently, the effect of a change in behavior of a small group of agents on the induced social identity cues $\tilde{\Pi}(\sigma, \Pi)$ will be larger.

Whether a '*Non-Neutral*' is unstable depends therefore on γ and the properties of the response function $\eta(\pi_x, p_x)$.

Minority Effect - There are two different ways in which agents can process π_x and p_x in their response function. They can either process the difference $\pi_x - p_x$ or the proportion $\frac{\pi_x}{p_x}$. When $p_0 = p_1$, this does not affect the local effects of a small change in a 'Neutral Regime'. This is nevertheless not true when $p_0 \neq p_1$. The effects of a change away from the '*Neutral Regime*' are still symmetric for both social groups when agents process the difference $\pi_x - p_x$, but not when agents process the proportion $\frac{\pi_x}{p_x}$ in their response function. Specifically, with the latter response function, the minority group will react more strongly to changes in social context than the majority group.

Degree of Asymmetry - The more strongly agents react to their social context in the belief formation process, the more social context will drive their choice behavior. Corollary 2 shows how this affects the degree of asymmetry we observe in a '*Non-Neutral Regime*'.

COROLLARY 2: *Take two response functions $\hat{\eta}$ and η, such that $\hat{\eta}(\pi, p) > \eta(\pi, p)$ for all $\pi > p$. Assume WLOG that a 'Non-Neutral Regime' exists in which $\pi^* > p$. Let π^*_η be the equilibrium value of π given a response function η. Then, $\pi^*_{\hat{\eta}} > \pi^*_\eta$.*

Welfare - If we consider a social planner maximizing the aggregate expected utility of all agents in the society, then a 'Non-Neutral Regime' is a Pareto improvement with respect to a 'Neutral Regime'. In a 'Neutral Regime', the strategies *Repress* and *Not Repress* are equivalent, and all agents form beliefs in the same way. In a 'Non-Neutral Regime', only those agents for whom it is strictly optimal will react to the asymmetries in social context. Therefore, in a 'Non-Neutral Regime', agents that learn to '*Not Repress*' are better off, while agents that learn to '*Repress*' are not made worse off. The Pareto optimality of a 'Non-Neutral Regime' should nevertheless be considered as a result that is driven by the simplifying assumptions made in the model, and can be easily contested. For example, a 'Non-Neutral Regime' is not necessarily Pareto optimal when beliefs have a direct effect on the probability of success, through for example confidence (Compte and Postlewaite, 2004), or when social context has a direct effect on someone's chances of success through some form of discrimination or stereotype threat (Steele, 2010). Finally, a 'Non-Neutral Regime' can become suboptimal when agents make systematic errors in learning their optimal strategy, when they do not correctly compute the long-term pay-offs of choosing a *Competitive* task, or when the strategy *Repress* becomes costly.

4 Social Identity Cues and Informational Policy

4.1 Data and Structure on Information

In this section, I shed light on how the data and structure agents put on this data affect the persistence of identity-driven choice behavior. Agents could for example process statistics stemming from who tries the *Competitive* task. It may take too long for a 'successful outcome' to be realized, or it may not be universally clear what a 'successful outcome' looks like. Agents may then believe that the fraction of people that try the *Competitive* task among those in their social group is indicative for their suitability to the task, or the degree of hostility in the environment (Chung, 2000). Agents could also observe those that try but fail.

The key driver of persistent asymmetry in social context is the population effect. Therefore, for a '*Non-Neutral Regime*' to exist, the structure agents put on their public information must capture this effect. This implies that a '*Non-Neutral Regime*' exists for most types of statistics, except for the within-group average success rates. To illustrate why a '*Non-Neutral Regime*' cannot exist when agents process this cue, consider an example in which the average success rate of women, π_1, is higher than the success rate of men, π_0. Corollary 1 shows that hence more women choose the *Competitive* task than men, but that this induces simultaneously a higher average success rate for men than for women. Consequently, the new social identity cues will induce more men to choose the *Competitive* task than women, which will induce a higher average success rate for women, and so forth.

This implies that we could eliminate the persistence of identity-driven choice behavior by influencing the statistics people take into account, and by making data that is often not available or hidden, such as those who tried but failed, more visible. The model also shows people react to their perception of social context, and predicts therefore that informational policy measures could be complementary to a real and maybe more costly change of social context, through for example affirmative action policy.

4.2 Misspecified Reaction Function

Example - Consider again the firm from the previous examples. Now, assume there are fewer women than men that have the qualifications to pursue a management career, but that agents hold an incorrect belief about the fraction of qualified women in the population. Specifically, let this incorrect belief be $\hat{p}_1 = \frac{1}{2}$, while the true fraction of qualified women is $p_1 < \frac{1}{2}$. Now, agents will incorrectly perceive a '*Neutral Regime*' as the same fraction of men and women in the pool of successful managers. Consequently, when '*Neutral Regime*' appears, agents will not interpret it as such and they will perceive women to be underrepresented, while men are perceived to be overrepresented. Because they now have a misspecified reaction function, women with $\alpha > \gamma$ will choose *Repress*, while men with $\alpha > \gamma$ will not, and vice versa. This drives the population towards a '*Non-Neutral Regime*' in which women will indeed be underrepresented among the successful managers. Corollary 3 formalizes this result.

COROLLARY 3: *Assume WLOG that agents hold a belief $\hat{p}_x > p_x$. Then there only exists a 'Non-Neutral Regime' in which $\pi_x < p_x$. A 'Neutral Regime' no longer exists.*

This shows how important it is to inform agents about the relevant fractions of social groups in the populations. If they cannot form correct beliefs about what a '*Neutral Regime*' looks like, it will never appear.

4.3 Individual Feedback

Because the options to manage confidence using social identity cues can be asymmetric across social types, similar types of individual feedback can have different effects on choice behavior across social groups. One could exploit these differences to boost diversity in educational and professional environments. For example, if students belonging to an ethnic majority are overrepresented in top educational pathways, only they have the option to boost up their beliefs. If one wants to induce more students belonging to ethnic minorities in these educational pathways, this could be achieved by giving those students, that have the capabilities and grades to succeed, systematically more positive feedback regarding their abilities. This would bias their individual-specific noisy perception upwards in a similar way as what majority students can achieve with the use of social context. These insights are in line with already existing programs that aim to enhance the confidence of underrepresented groups to increase diversity.

5 Discussion

5.1 Imperfect Learning

I assume agents are perfectly able to learn their individually optimal strategies. The main objective of this assumption is to show that, even when agents learn perfectly, asymmetries in choice behavior can persist. In this section, I discuss what happens when this learning is imperfect.

The equilibrium model can be adjusted to allow for imperfect learning as follows. The induced probability to choose the *Competitive* task for an agent of type

$\{\alpha, \theta\}$ in a social context Π, playing strategy σ_i as presented in Equation (6) can be written as,

$$\Phi_{\alpha,x,\sigma_i,\Pi} = \Sigma_{\sigma \in \{R,NR\}} P(\sigma_i = \sigma|\alpha, x, \Pi) P(\hat{p}_i^\sigma > \gamma|\alpha) \tag{13}$$

In the case of perfect learning, $P(\sigma_i = \sigma|\alpha, x, \Pi) \in \{0, 1\}$, while in the case of imperfect learning, $P(\sigma_i = \sigma|\alpha, x, \Pi) \in [0, 1]$. Let λ be an exogenous learning process. Then, any such learning process implies a probability $P^\lambda(\sigma_i = \sigma|\alpha, x, \Pi)$. As long as $P^\lambda(\sigma_i = \sigma^*|\alpha, x, \Pi) > P^\lambda(\sigma_i \neq \sigma^*|\alpha, x, \Pi)$, we observe differences in $\Phi_{\alpha,x,\sigma_i,\Pi}$ across agents with a different characteristic θ. This will induce differences in choice behavior and further reasoning continues along the same lines as in a model with perfect learning.

Finally, if failing to learn means failing to *Repress* the influence of social context when this is optimal, imperfect learning implies that more agents than optimal use their social identity cue in belief formation. In this case, imperfect learning would increase the strength of the population effect. If failing to learn implies that agents make random mistakes, imperfect learning may decrease the strength of the population effect.

5.2 Towards a More Realistic Model

To simplify the model, I assume α and γ are fixed over an agent's lifetime. If agents can learn their optimal strategy conditional on α, this may raise the question why they are not able to retrieve α itself, and hence, their optimal task. Indeed, a sufficiently sophisticated agent could interpret σ_i^* as an extra signal regarding α_i. When α and γ are fixed, this would be a perfect signal. Therefore, agents that understand the structure of the model, could use σ_i^* to learn their optimal choice of task. Consequently, social context has no differential effect on choice behavior anymore across social groups, and a '*Non-Neutral Regime*' cannot exist. This reasoning is nevertheless too much driven by the simplifying assumption of keeping α and γ fixed.

Consider now a more realistic model, in which the values of α and γ vary over the lifetime of an agent, such that agents learn from a series of related, but slightly different tasks. The optimal strategy will be conditional on whether on average during the learning process α has been above or below γ. In a particular choice context, we will therefore have a fraction of agents with $\alpha > \gamma$ belonging to the socially more successful group that will have learned to *Not Repress*, but also a fraction that will have learned to *Repress*. Consequently, σ^* becomes an imperfect signal regarding α and sophisticated agents will no longer be able to retrieve their true α. They may learn that on average α is below or above γ, and derive from that a strategy to always choose the *Competitive* or *Non-Competitive* task. As long as $\hat{\alpha}$ is informative about α, the belief formation rule following from σ^* will nevertheless outperform such a strategy.[14] Even more sophisticated agents may want to use this extra signal to further improve upon their decision, which eventually means forging a third belief: $\hat{p}_i^I$, the resulting belief from the inference process I. We can therefore account for this type of sophistication in the model by enriching the set of strategies[15], and consider agents that compare three possible beliefs; $\hat{p}_i^R, \hat{p}_i^{NR}$ or $\hat{p}_i^I$.

[14]The same reasoning applies to agents that can only imperfectly learn their individually optimal strategy in the simplified model.

[15]In the spirit of Compte and Postlewaite (2019), this would limit the degree to which agents can compare these strategies, since one cannot compare more strategies without at the same time altering the accuracy with which one can compare them.

To analyze the effects of this enriched set of strategies on the equilibrium results, consider again the example of the firm. If the belief $\hat{p}_i^I$ results from a correct inference process, then a woman learning to *Not Repress*, will be able to make the inference that, on average, she is good enough to choose leadership-related tasks, while a woman learning to *Repress*, will make the inference that, on average, she is not good enough to choose leadership-related tasks. The opposite applies to men. Now, the key aspect driving the existence of a 'Non-Neutral Regime', namely the fact that belief formation is type-contingent, disappears. One could argue nevertheless that the latter inference seems more difficult to make than the former, since it requires a more elaborate thinking. It seems easier for a woman to infer that, when she believes it is relevant that women are overrepresented among the successful individuals, she is also likely to succeed. But it is much less straightforward for a woman to infer that, if it is not optimal to use the female-driven bias, then it must be that she has low chances of success. She may instead conclude the statistic is not providing useful or relevant information regarding her own abilities. Similarly, when a man learns to *Not Repress*, it may be easy for him to infer that, like all other men, his chances of success are not that great either. It is much less straightforward for him to infer that, if *Repress* is the better strategy, then he must be good. Therefore, if the belief $\hat{p}_i^I$ follows from a correct and complete inference process, a 'Non-Neutral Regime' can no longer exist. On the other hand, a partial inference process, like the one described above, would exacerbate the phenomenon.

5.3 Discrimination and Other Effects of Social Identity

To isolate the mechanism I want to describe in this paper, I assume α and θ are uncorrelated. In reality, there may be other effects of social identity on choice. I categorize the effects described in the current literature as either direct effects of social identity on utility derived from for example self-image or punishment by others, or direct effects on the agent's real or perception of her chances of success in each type of task because of for example discrimination or the anticipation of possible discrimination. In this section, I aim to show with a simple model how these effects would interact with the mechanism presented in this paper.

Let $\tau \in [0, 1)$ be a tax agents pay when choosing an action $a \in \{C, NC\}$ for which their social group is underrepresented. In other words, $\tau > 0$ for an agent with characteristic x, when $a = C$ and $\pi_x < p_x$ or when $a = NC$ and $\pi_x > p_x$. Otherwise, $\tau = 0$. The tax τ affects the subjective expected utility of undertaking an action that agents compare. To illustrate, assume $\pi_1 > p_1$ and $\pi_0 < p_0$. To choose their action a, agents with $\theta = 1$ will now compare the subjective expected utility of undertaking the *Competitive* task, $\hat{p}_i^{\sigma^*}$, with the subjective expected utility of undertaking the *Non-Competitive* task, $\gamma[1 - \tau]$, while agents with $\theta = 0$ will compare $\hat{p}_i^{\sigma^*}[1 - \tau]$ to γ. Here, τ can be interpreted both as a negative effect on utility, and as a negative effect on the probability of success in the respective task, or the agent's perception thereof.

Agents with $\theta = 1$ will choose a threshold $\gamma_i \in \{\gamma[1 - \tau], \frac{\gamma[1-\tau]}{\eta_1}\}$, while agents with $\theta = 0$ will choose a threshold $\gamma_i \in \{\frac{\gamma}{[1-\tau]}, \frac{\gamma}{[1-\tau]\eta_0}\}$. Therefore, a tax $\tau > 0$ moves the set of thresholds agents with $\theta = 1$ compare downwards, making it more likely they choose the *Competitive* task independent of their value of α, while it moves the set of thresholds agents with $\theta = 0$ compare upwards, making it more likely they choose the *Non-Competitive* task. Furthermore, we can write the induced number of successful agents in the *Competitive* task with $\theta = 1$ as follows,

$$S_1 = p_1 \left[\int_{\hat{\alpha} > \frac{\gamma}{\eta_1}} g_\alpha(\hat{\alpha}) d\hat{\alpha} + \int_{\frac{\gamma[1-\tau]}{\eta_1} < \hat{\alpha} < \frac{\gamma}{\eta_1}} g_\alpha(\hat{\alpha}) d\hat{\alpha} \right] \int_{\alpha > \gamma} \alpha f(\alpha) d\alpha + \tag{14}$$

$$p_1 \left[\int_{\hat{\alpha} > \gamma} g_\alpha(\hat{\alpha}) d\hat{\alpha} + \int_{\gamma[1-\tau] < \hat{\alpha} < \gamma} g_\alpha(\hat{\alpha}) d\hat{\alpha} \right] \int_{\alpha < \gamma} \alpha f(\alpha) d\alpha \tag{15}$$

while the number of successful agents in the *Competitive* task with $\theta = 0$ is given by,

$$S_0 = p_0 \left[\int_{\hat{\alpha} > \gamma} g_\alpha(\hat{\alpha}) d\hat{\alpha} - \int_{\gamma < \hat{\alpha} < \frac{\gamma}{[1-\tau]}} g_\alpha(\hat{\alpha}) d\hat{\alpha} \right] \int_{\alpha > \gamma} \alpha f(\alpha) d\alpha + \tag{16}$$

$$p_0 \left[\int_{\hat{\alpha} > \frac{\gamma}{\eta_0}} g_\alpha(\hat{\alpha}) d\hat{\alpha} - \int_{\frac{\gamma}{\eta_0} < \hat{\alpha} < \frac{\gamma}{[1-\tau]\eta_0}} g_\alpha(\hat{\alpha}) d\hat{\alpha} \right] \int_{\alpha < \gamma} \alpha f(\alpha) d\alpha \tag{17}$$

In these equations, I separate the effects of social cues, captured by each first term in the brackets, and the effects of the tax τ, captured by each second term in the brackets. The first equation shows how S_1 increases through agents with $\alpha > \gamma$ and $\frac{\gamma[1-\tau]}{\eta_1} < \hat{\alpha} < \frac{\gamma}{\eta_1}$, and agents with $\alpha < \gamma$ and $\frac{\gamma}{\eta_0} < \hat{\alpha} < \frac{\gamma}{[1-\tau]\eta_0}$, that choose $a = C$ solely driven by the effect of τ. On the other hand, S_0 decreases through agents with $\alpha > \gamma$ and $\gamma < \hat{\alpha} < \frac{\gamma}{[1-\tau]}$, and agents with $\alpha < \gamma$ and $\frac{\gamma}{\eta_0} < \hat{\alpha} < \frac{\gamma}{[1-\tau]\eta_0}$, that in the absence of the effect of τ would have chosen $a = C$, but now choose $a = NC$.

As τ goes towards 1, we go towards an extreme case of the model in which all agents with $\theta = 1$ choose the *Competitive* task and all agents with $\theta = 0$ choose the *Non-Competitive* task. On the other hand, as τ goes towards zero, we move towards the case discussed in this paper. This simple analysis provides therefore the intuition for how the results presented in this paper are robust in a setting in which we introduce other effects of social identity on decision making, and shows how the mechanism presented in this paper and other effects of social identity on choice reinforce each other.

6 Social Identity and Belief Formation

Social identity, belief formation and choice behavior are topics that are widely studied outside economics. The idea that social context and identity affect a person's perception of her own abilities finds its origin in the field of social psychology. Hogg and Grieve (1999) discusses how in the process of depersonalization, which is associated with social identification, individual and concomitant unshared beliefs, attitudes, feelings, and behaviors are replaced by an in-group prototype that prescribes shared beliefs, attitudes, feelings and behaviors. Similarly, Seligman (2006) shows how people can interpret numerous failures from others like them as evidence that they will fail as well. Finally, Steele (2010) discusses how the psyche of the individual gets damaged by bad images of their group projected in society. Repeated exposure to these images causes them to be internalized, leading to low self esteem, low expectations, low motivation and self doubt.

There is also a large literature showing how social identity affects choice behavior. For example, Smith et al. (2007) shows that, when people complete a high stereotype-threat test, they report decreased task interest. Davies et al. (2002) argues how the combination of decreased enjoyment and diminished self-confidence explains why women experiencing stereotype threat report less interest in math and science fields and weaker leadership aspirations compared to men or non-threatened women. Similarly, Banaji and Greenwald (2016) shows how implicit associations picked up from social context by our automatic brain affect our behavior, such as the intellectual pursuits we select, and Perry et al. (2003) discusses how people tend to protect themselves against stereotype threat by ceasing to care about the domain in which the stereotype applies. Finally, Oh (2023) shows how Indian workers are willing to forego substantial payments to avoid tasks that are associated with other castes.

Hogg and Grieve (1999) defines two classes of motivation for social identification. The first motivation is self-esteem. People are motivated to maintain or achieve positive distinctiveness for their own group relative to other groups, because intergroup evaluation is self-evaluation. This idea is introduced in economics by Akerlof (2016). The second motivation is subjective uncertainty reduction. Subjective certainty gives people confidence about how to behave, and what to expect from their physical and social environment. This is related to Atkin et al. (2021) and Shayo (2009), that show how ethnic and religious identities are determined by group status, group salience and the market cost of following a group's prescribed behaviors. Benabou and Tirole (2011) introduces an idea very much related to this paper, namely that identity investments are driven by welfare maximization considerations.

Finally, it is not clear whether people are aware of their social identification and its effects on their behavior. Purdie-Vaughns et al. (2008) and Marx and Goff (2005) show that black professionals and students are often aware of the presence of stereotype threat, and Steele et al. (2002) shows that some female undergraduates report in a math and science report that they believe they have weak abilities because of their gender. At the same time, Stone et al. (1999) and Leyens et al. (2000) show that white athletes and men fail to report anxiety when they experience stereotype threat. Banaji and Greenwald (2016) and Murden (2020) argue that the effects of social context on behavior are largely determined by the automatic part of our brain, outside of our awareness. In the model, I make no assumptions regarding whether people are aware of their choice of belief formation process, and the model can be consistent with both scenarios.

7 Conclusion

This paper shows how people can use statistics about the prevalence of their social group among the successful individuals in a task to cope with the negative effects of momentary noise decision making. Although this behavior is optimal from an individual perspective, it can create persistent differences in choice behavior across a priori identical social groups. If we want to eliminate asymmetries across social groups, taking care of discrimination, skill-differences or social pressure is therefore not enough. I discuss implications for informational policy to address this.

The insights provided by this paper point to several areas for future research. First, an important limitation of this model is that it assumes homogeneity in both the information agents retrieve from social context and in the way in which agents process this information. Social networks may nevertheless play an important role in the perception of the social environment. This could create heterogeneity in social cues that may be correlated with observable characteristics through variables such as income, neighborhood or education. Secondly, in the paper agents are perfectly able to learn their optimal strategies. The process in which agents learn how to interpret social context is nevertheless an interesting research topic on its own. Social context could influence this learning process directly, through for example stigmatization, discrimination, implicit biases and expectations, social pressures or stereotype threat, which could induce learning traps that could be asymmetric across social groups. Psychological factors may also play a role, such as the shame to learn you are not good enough to undertake a task, even though you belong to the socially more successful group. These factors may affect the set of belief formation heuristics people choose from. A deeper understanding of these issues would allow us to better make the step from the theoretical framework to the real world, and derive more concrete policy implications.

References

G.A. Akerlof and R.E. Kranton. Economics and identity. *Quarterly Journal of Economics*, 115(3):715–753, 2000.

R. Akerlof. We thinking and its consequences. *American Economic Review*, 106 (5):415–419, 2016.

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.

M.R. Banaji and A.G. Greenwald. *Blindspot*. Random House Publishing Group, 2016.

A.V. Banerjee. A simple model of herd behavior. *The Quarterly Journal of Economics*, 107(3):797–817, 08 1992.

H. Battu, M. McDonald, and Y. Zenou. Oppositional identities and the labor market. *Journal of Population Economics*, 20:643–667, 2007.

R. Benabou and J. Tirole. Self-confidence and personal motivation. *The Quarterly Journal of Economics*, 117(3):871–915, 08 2002.

R. Benabou and J. Tirole. Incentives and prosocial behavior. *American Economic Review*, 96(5):1652–1678, 2006.

R. Benabou and J. Tirole. Identity, morals and taboos: Beliefs as assets. *The Quarterly Journal of Economics*, 126(2):805–855, 05 2011.

J. Benhabib, Alberto Bisin, and M. Jackson. *Handbook of Social Economics*, chapter 5, pages 201–213. North-Holland, 2010. Theories fo Statistical Discrimination and Affirmative Action: A Survey.

F.D. Blau and L.M. Kahn. The gender wage gap: Extent, trends, and explanations. *Journal of Economic Literature*, 55(3):789–865, 2017.

P. Bordalo, N. Gennaioli, and A. Shleifer. Stereotypes. *Quarterly Journal of Economics*, 131(4):1753–1794, 2016.

M.K. Brunnermeier and J.A. Parker. Optimal expectations. *The American Economic Review*, 95(4):1092–1118, 2005.

L. Bursztyn and R. Jensen. Social image and economic behavior in the field: Identifying, understanding and shaping social pressure. *Annual Review of Economics*, 9:131–153, 2017.

T. Buser, M. Niederle, and H. Oosterbeek. Gender, competitiveness and career choices. *The Quarterly Journal of Economics*, 129(3):1409–1447, 2014.

J.P. Carvalho and B. Pradelski. Identity and underrepresenation: Interactions between race and gender. *Journal of Public Economics*, 216:104764, 2022.

K. Chung. Role models and arguments for affirmative action. *American Economic Review*, 90(3):640–648, 2000.

S. Coate and G.C. Loury. Will affirmative-action policies eliminate negative stereotypes? *The American Economic Review*, 83(5):1220–1240, 1993.

K.B. Coffman. Evidence on self-stereotyping and the contribution of ideas. *The Quarterly Journal of Economics*, 129(4):1625–1660, 2014.

C. Compte and A. Postlewaite. *Ignorance and Uncertainty*. Cambridge University Press, 2019.

O. Compte and A. Postlewaite. Confidence-enhanced performance. *American Economic Review*, 94(5):1536–1557, 2004.

P.G. Davies, S.J. Spencer, D.M. Quinn, and R. Gerhardstein. Consuming images: how television commercials that elicit stereotype threat can restrain women academically and professionally. *Personal. Soc. Psychol. Bull.*, 28(12):1615–1628, 2002.

E. Dekel, J.C. Ely, and O. Yilankaya. Evolution of preferences. *Review of Economic Studies*, 74:685–704, July 2007.

J. Elster. Rationality and economics. *The Economic Journal*, 106(438):1386–1397, 1996.

I. Esponda and M. Pouzo. Berk-nash equilibrium: A framework for modeling agents with misspecified models. *Econometrica*, 84(3):1093–1130, 2016.

C.L. Exley and J.B. Kessler. The gender gap in self-promotion. *Quarterly Journal of Economics*, 137(3):1345–1381, 2022.

K. Fiedler and H. Bless. *Emotions and belief: How feelings influence thoughts*, chapter The formation of beliefs at the interface of affective and cognitive processes, pages 144–170. Cambridge University Press, 2000.

J.A. Flory, Leibbrandt A., and J.A. List. Do competitive work places deter female workers? a large-scale natural field experiment on gender differences in job-entry decisions. *The Review of Economic Studies*, 82(1):122–155, 2015.

M. Frick, R. Iijima, and Y. Ishii. Dispersed behavior and perceptions in assortative societies. Discussion Paper 2650, Cowles Foundation, 2018.

D. Fudenberg and D.K. Levine. Self-confirming equilibrium. *Econometrica*, 61(3): 523–545, 1993.

N. Guyon and E. Huillery. Biased aspirations and social inequality at school: Evidence from french teenagers. *The Economic Journal*, 131(634):745–796, 2021.

S. Hodge, J.W. Burden, L.E. Robinson, and R. Bennett. Theorizing on the stereotyping of black male student-athletes: Issues and implications. *Journal for the Study of Sports and Athletes in Education*, 2(2):203–226, 2008.

K. Hoff and J.E. Stiglitz. Equilibrium fictions: A cognitive approach to societal rigidity. *American Economic Review: Papers and Proceedings*, 100:141–146, 2010.

K. Hoff and J.E. Stiglitz. Striving for balance in economics: Towards a theory of the social determination of behavior. *Journal of Economics Behavior and Organization*, 126:25–57, 2016.

M.A. Hogg and P. Grieve. Social identity theory and the crisis of confidence in social psychology: A commentary, and some research on uncertainty reduction. *Asian Journal of Social Psychology*, 2:79–93, 1999.

J.P. Leyens, M. Desert, J.C. Croizet, and C. Darcis. Stereotype threat: Are lower status and history of stigmatization preconditions of stereotype threat? *Personal. Soc. Psychol. Bull.*, 26(10):1189–1199, 2000.

Q. Lippmann and C. Senik. Math, girls and socialism. *Journal of Comparative Economics*, 46(3):874–888, 2018.

C.W. Liqui-Lung. Intersectionality and individual decision making. Working paper, University of Cambridge, 2023.

D.M. Marx and P.A. Goff. Clearing the air: the effect of experimenter race on targets test performance and subjective experience. *Br. J. Soc. Psychol.*, 44(4): 645–657, 2005.

P. Mukherjee. The effects of social identity on aspirations and learning outcomes: A field experiment in rural india. Working paper, International Growth Center, 2017.

F. Murden. *Mirror Thinking.* Bloomsbury Sigma, 2020.

M. Niederle and L. Vesterlund. Do women shy away from competition? do men compete too much? *The Quarterly Journal of Economics*, 122(3):1067–1101, 2007.

S. Oh. Does identity affect labor supply. *American Economic Review*, 113(8): 2055–2083, 2023.

P. Onuchic. Recent contributions to theories of discrimination. Working paper, Oxford University, 2023.

T. Perry, C. Steele, and A.G. Hilliard III. *Young, Gifted, and Black*. Beacon Press, 2003.

M. Peski and B. Szentes. Spontaneous discrimination. *American Economic Review*, 103(6):2412–2436, 2013.

S. Peters. *A Path through the Jungle*. Mindfield Media, 2021.

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

T. Piketty. Social mobility and redistributive politics. *The Quarterly Journal of Economics*, 110(3):551–584, 1995.

E. Pronin, C.M. Steel, and L. Ross. Identity bifurcation in response to stereotype threat: women and mathematics. *Journal of Experimental Social Psychology*, 40(2):152–168, 2004.

V. Purdie-Vaughns, C.M. Steel, P.G. Davies, R. Ditlmann, and J.R. Crosby. Social identity contingencies: how diversity cues signal threat or safety for african americans in mainstream institutions. *Journal of Personal. Social Psychology*, 94(4):615–630, 2008.

M. Seligman. *Learned Optimism: How to change your mind and your life*. Random House USA Inc, 2006.

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.

J.L. Smith, C. Sansone, and P.H. White. The stereotyped task engagement process: the role of interest and achievement motivation. *Journal of Education and Psychology*, 99(1):99–114, 2007.

R. Spiegler. Bayesian networks and boundedly rational expectations. *The Quarterly Journal of Economics*, 131(3):1243–1290, 2016.

C.M. Steele. *Whistling Vivaldi: How stereotypes affect us and what we can do.* W.W. Norton and Company, Inc, 2010.

J. Steele, J.B. James, and R.C. Barnett. Learning in a mans world: examining the perceptions of undergraduate women in male-dominated academic areas. *Psychol. Women Q.*, 26(1):46–50, 2002.

J. Stone, C.I. Lynch, M. Sjomeling, and J.M. Darley. Stereotype threat effects on black and white athletic performance. *J. Personal. Soc. Psychol.*, 77(6): 1213–1228, 1999.

A. Wolitzky. Learning from others' outcomes. *American Economic Review*, 108 (10):2763–2801, 2018.

A Mathematical Appendix

PROPOSITION 1 (Individually Optimal Belief Formation): *The individually optimal strategies σ^* given an agent's type $\{\alpha, \theta\}$ are the following:*

- *The individually optimal strategy σ^* is 'Not Repress' for agents of type $\{\alpha, x\}$ such that $\alpha > \gamma$ and $\pi_x > p_x$ or $\alpha < \gamma$ and $\pi_x < p_x$*
- *The individually optimal strategy σ^* is 'Repress' for agents of type $\{\alpha, x\}$ such that $\alpha > \gamma$ and $\pi_x < p_x$ or $\alpha < \gamma$ and $\pi_x > p_x$*

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 to take action $a = C$. Therefore, V_i is larger when playing NR than when playing R if and only if $\Phi_{\alpha,x,NR,\Pi} \geq \Phi_{\alpha,x,R,\Pi}$. Since $\Phi_{\alpha,x,\sigma,\Pi} = P(\hat{\alpha} > \gamma^{\sigma}|\alpha)$, this is the case when $\gamma^{NR} < \gamma^{R}$. This is true if and only if $\pi_x \geq p_x$. Therefore, NR is only an optimal strategy for agents with $\alpha > \gamma$ when their observable characteristic $\theta_i = x$ is such that the social identity cue $\pi_x \geq p_x$. If this is not the case, they are better off choosing strategy $\sigma_i = R$. Vice versa for agents with $\alpha < \gamma$, the welfare-maximizing choice is to take action $a = NC$. Therefore, V_i i larger when playing NR than when playing R if and only if $\Phi_{\alpha,x,NR,\Pi} \leq \Phi_{\alpha,x,R,\Pi}$. This is the case if and only if $\gamma^{NR} > \gamma^{R}$, meaning that we need $\pi_x \leq p_x$. Therefore, agents with α should only choose strategy $\sigma_i = NR$, when their observable characteristic is such that $\pi_x \leq p_x$. Otherwise, they are better off choosing strategy $\sigma_i = R$. ■

LEMMA 1: *If there exist $\underline{\pi} > p$ and x such that $\tilde{\pi_x}(\sigma^*, \underline{\pi}) > \underline{\pi}$, there exists a Non-Neutral Regime such that $\pi_x^* \neq p_x$.*

Proof. We know,

$$\tilde{\pi}_x(\sigma^*,\pi) = \int_{\alpha>\gamma} \alpha P\left(\hat{\alpha} > \frac{\gamma}{\eta(\pi,p)}\right) f(\hat{\alpha}|\alpha)d\hat{\alpha}d\alpha + \int_{\alpha<\gamma} \alpha P(\hat{\alpha} > \gamma) f(\hat{\alpha}|\alpha)d\hat{\alpha}d\alpha \tag{18}$$

When $\eta(\pi,p)$ is monotonic, if $\pi' > \pi$, then $\eta(\pi',p) > \eta(\pi,p)$, and $\tilde{\pi}_x(\sigma^*,\pi)$ is continuous in π.[16] Furthermore, from Equation (18) it follows $\tilde{\pi}_x(\sigma^*,\pi)$ is increasing in π. Then, when there exist $\underline{\pi} > p$ and x such that $\tilde{\pi}_x(\sigma^*,\underline{\pi}) > \underline{\pi}$, then $\forall \pi > \underline{\pi}$, we have $\tilde{\pi}_x(\sigma^*,\pi) > \underline{\pi}$. In Section 3.2, I showed how π is bounded from above by an upperbound $\overline{\pi}$. Consequently, $\tilde{\pi}_x(\sigma,\pi)$ is continuous in π on the closed set $[\underline{\pi},\overline{\pi}]$, and following Brouwer's fixed point theorem, there exists a fixed point π^*.

■

PROPOSITION 2 (Existence Non-Neutral Regime): *Let $\delta > 0$ be a small value with which we disturb a 'Neutral Regime'. When $p_1 = p_0$, a sufficient condition for the co-existence of a stable 'Non Neutral Regime' with a 'Neutral Regime' is as follows,*

$$\frac{\partial \gamma^{NR}}{\partial \delta} \frac{|\frac{\partial S}{\partial \gamma}|}{S} > 4 \tag{19}$$

where $\gamma^{NR} = \frac{\gamma}{\eta_x}$ and $S = \int \alpha G_\alpha(\gamma) f(\alpha) d\alpha$.

Proof. When the condition of Lemma 1 holds, a 'Non-Neutral Regime' co-exists with a 'Neutral Regime'. We can show that this is the case when either there is a

[16] Even without continuity, a monotonic function has a finite number of jumps. Because α is continuous, $\tilde{\pi}(\pi)$ is continuous.

jump in $\eta(\pi, p)$ at $\pi = p$, or when

$$\frac{\partial \tilde{\pi}_x(\sigma^*, \pi)}{\partial \pi}|_{\pi=p} > 1 \tag{20}$$

In the following, I derive a sufficient condition for the case $p_1 = p_0$. Consider a slight perturbation of a 'Neutral Regime', such that $\pi_1 = p_1 + \delta$, while $\pi_0 = p_0 - \delta$. Assume WLOG that the response function is continuous and such that agents process the difference $\pi_x - p_x$[17]. Then, $\eta_1 = \eta(\pi_1 + \delta - p_1) = \eta(0) + \eta'(0)[\pi_1 + \delta - p_1]$. Therefore,

$$\frac{\gamma}{\eta_1} \simeq \frac{\gamma}{\eta(0)}[1 - \frac{\eta'(0)}{\eta(0)}\delta] \tag{21}$$

Because $\eta(0) = 1$, $\gamma - \frac{\gamma}{\eta_1} \simeq \gamma\eta'(0)\delta$. This shows that this change is multiplicative in γ. Furthermore, let $\Delta_x = \gamma - \frac{\gamma}{\eta_x}$. Then, there is a symmetry, such that $\Delta_1 = -\Delta_0$. Let $S_1 = p_1 \int \alpha G_\alpha(\gamma) f(\alpha) d\alpha$ and $S_0 = p_0 \int \alpha G_\alpha(\gamma) f(\alpha) d\alpha$, and $S'_x = \frac{\partial S_x}{\partial \delta}$. Then,

$$S'_1 = S_1 + p_1 \int_{\alpha > \gamma} \alpha f(\alpha) d\alpha \int_{\gamma - \Delta_1 < \hat{\alpha} < \gamma} g_\alpha(\hat{\alpha}) d\hat{\alpha} \tag{22}$$

where $\int_{\gamma - \Delta_1 < \hat{\alpha} < \gamma} g_\alpha(\hat{\alpha}) d\hat{\alpha} \approx g_\alpha(\gamma)\Delta_1$. Similar for S'_0. For Δ_θ arbitrarily small,

$$\frac{S'_1}{S'_0} = \frac{p_1}{p_0}\left[1 + \Delta \frac{|\frac{\partial S}{\partial \gamma}|}{S}\right] \tag{23}$$

[17] When agents process the proportion $\frac{\pi_x}{p_x}$, $\Delta = \delta\eta'(0)\gamma$, and locally we have $\eta\left(\frac{\pi_x}{p_x}\right) \approx \eta(1 + 2(\pi_x - p_x))$. Therefore, in the case $p_0 = p_1$, the sufficient condition in (27) also applies.

with $\Delta = \gamma\eta'(0)\delta$. A 'Neutral Regime' becomes unstable when,

$$\frac{p_1}{p_0}\left[1 + \Delta\frac{|\frac{\partial S}{\partial \gamma}|}{S}\right] > \frac{p_1 + \delta}{p_0 - \delta} \tag{24}$$

When $p_1 = p_0 = \frac{1}{2}$, this is the case when

$$\gamma\eta'(0)\frac{|\frac{\partial S}{\partial \gamma}|}{S} > 4 \tag{25}$$

Finally, we note that $\gamma\eta'(0) = \frac{\partial(\gamma - \frac{\gamma}{\eta x})}{\partial \delta} = \frac{\partial \gamma^{NR}}{\partial \delta}$.

■

COROLLARY 1: *Let $x' \in \{0, 1\}$ be the complement of x and assume WLOG that $\pi_x > \pi_{x'}$. The optimal use of social identity cues has both a population effect, such that $\Phi_{\alpha,x,\sigma_i,\Pi} > \Phi_{\alpha,x',\sigma_i,\Pi}$ and a selection effect, such that $E(\alpha|a = C, x) < E(\alpha|a = C, x')$. The strength of both effects is such that the order $\pi_x > \pi_{x'}$ will always be preserved.*

Proof. Assume WLOG that $\pi_1 > \pi_0$. Then, $\frac{\gamma}{\eta_1} < \gamma$, while $\frac{\gamma}{\eta_0} > \gamma$. Therefore, all agents with $\alpha > \gamma$ and $\theta_i = 1$ will choose $\gamma_i = \frac{\gamma}{\eta_1}$, while all agents with $\alpha > \gamma$ and $\theta_i = 0$ will choose $\gamma_i = \gamma$. On the other hand, agents with $\alpha < \gamma$ and $\theta_i = 1$ will choose $\gamma_i = \gamma$, while similar agents with $\theta_i = 0$ will choose $\gamma_i = \frac{\gamma}{\eta_0}$. In both cases, the threshold agents with $\theta_i = 1$ choose is lower than the threshold agents with $\theta_i = 0$ choose. Consequently, $\Phi_{\alpha,1,NR,\Pi} > \Phi_{\alpha,0,R,\Pi}$ for all α. At the same times, this implies agents with $\theta_i = 1$ will choose the *Competitive* task for on average higher realizations of $\hat{\alpha}$. Because $E(\hat{\alpha}) = \alpha$, these agents will on average also have higher true ability levels α, which leads to the selection effect

$E(\alpha|a=C,1) < E(\alpha|a=C,0)$. Finally, we can show that,

$$\frac{\partial\tilde{\pi}_1(\pi_1,\sigma)}{\partial\pi_1} = \frac{\pi_1\left[\frac{\partial S_1}{\partial\pi_1} - \frac{\partial S_0}{\partial\pi_1}\right]}{S} \tag{26}$$

where $S_1 = p_1 \int \alpha\Phi_{\alpha,1,\sigma,\Pi} f(\alpha)d\alpha$ and $S_0 = p_0 \int \alpha\Phi_{\alpha,0,\sigma,\Pi} f(\alpha)d\alpha$ denote the number of successful agents at the Competitive task with respectively $\theta_i = 1$ and $\theta_i = 0$. Furthermore, $\frac{\partial S_1}{\partial\pi_1}$ and $\frac{\partial S_0}{\partial\pi_1}$ for $\pi_1 \in [p_1, 1]$ are given by,

$$\frac{\partial S_1}{\partial\pi_1} = p_1 \int_{\alpha>\gamma} \alpha g_\alpha\left(\frac{\gamma}{\eta_1}\right)\frac{\gamma}{\eta_1^2}\frac{\partial\eta(\pi_1,p_1)}{\partial\pi_1} f(\alpha)d\alpha$$
$$\frac{\partial S_0}{\partial\pi_1} = -p_0 \int_{\alpha<\gamma} \alpha g_\alpha\left(\frac{\gamma}{\eta_0}\right)\frac{\gamma}{\eta_0^2}\frac{\partial\eta(\pi_1,p_1)}{\partial\pi_1} f(\alpha)d\alpha$$

Therefore, $\frac{\partial S_1}{\partial\pi_1} > 0$, while $\frac{\partial S_0}{\partial\pi_1} < 0$. Therefore, $\frac{\partial\tilde{\pi}_1(\pi_1,\sigma)}{\partial\pi_1} > 0$ and the selection and population effect will not reverse the order $\pi_1 > \pi_0$.

■

COROLLARY 2: *Take two response functions $\hat{\eta}$ and η, such that $\hat{\eta}(\pi,p) > \eta(\pi,p)$ for all $\pi > p$. Assume WLOG that a 'Non-Neutral Regime' exists in which $\pi^* > p$. Let π^*_η be the equilibrium value of π given a response function η. Then, $\pi^*_{\hat{\eta}} > \pi^*_\eta$.*

Proof. Let $\eta(\pi,p)$ be a response function such that, given γ, the condition of Lemma 1 holds. Then, a 'Non-Neutral Regime' also exists for any response function $\hat{\eta}(\pi,p)$, such that $\hat{\eta}(\pi,p) > \eta(\pi,p)$. Let $\tilde{\pi}_{\eta,x}(\pi,\sigma)$ be the induced value of π for a response function η. Then, if $\hat{\eta}(\pi,p) > \eta(\pi,p)$ for all $\pi > p$,

$$\tilde{\pi}_{\hat{\eta},x}(\pi,\sigma) > \tilde{\pi}_{\eta,x}(\pi,\sigma) \quad \forall\pi > p \tag{27}$$

Consequently, let π^*_η be the equilibrium value of π that arises in a 'Non-Neutral Regime' for a response function η. Then,

$$\pi^{(1)} \equiv \tilde{\pi}_{\hat{\eta},x}(\pi^*_\eta, \sigma) > \tilde{\pi}_{\eta,x}(\pi^*_\eta, \sigma) = \pi^*_\eta \tag{28}$$

which implies that,

$$\pi^{(2)} \equiv \tilde{\pi}_{\hat{\eta},x}(\pi^{(1)}, \sigma) > \tilde{\pi}_{\hat{\eta},x}(\pi^*_\eta, \sigma) \equiv \pi^{(1)} \tag{29}$$

and,

$$\pi^{(3)} \equiv \tilde{\pi}_{\hat{\eta},x}(\pi^{(2)}, \sigma) > \tilde{\pi}_{\hat{\eta},1}(\pi^{(1)}, \sigma) \equiv \pi^{(2)} \tag{30}$$

This sequence converges to $\pi^*_{\hat{\eta}} = \tilde{\pi}_{\hat{\eta},x}(\pi^*_{\hat{\eta}}, \sigma)$ and is everywhere above π^*_η and below the upper bound $\overline{\pi}$ on π. This shows that, for any response function $\hat{\eta}(\pi, p)$ such that $\hat{\eta}(\pi, p) > \eta(\pi, p)$ for all $\pi > p$, in equilibrium

$$\pi^*_{\hat{\eta}} > \pi^*_\eta \tag{31}$$

■

COROLLARY 3: *Assume WLOG that agents hold a belief $\hat{p}_x > p_x$. Then there only exists a 'Non-Neutral Regime' in which $\pi_x < p_x$. A 'Neutral Regime' no longer exists.*

Proof. Assume WLOG that $\hat{p}_0 > p_0$. This means that,

$$\eta(\pi_0, p_0) = \begin{cases} > 1 & \text{if } \pi_0 > \hat{p}_0 \\ 1 & \text{if } \pi_0 = \hat{p}_0 \\ < 1 & \text{if } \pi_0 < \hat{p}_0 \end{cases} \tag{32}$$

and consequently, when $\pi_0 = p_0$, $\eta_0 < 1$. This implies that $\tilde{\pi}_0(p_0, \sigma) < p_0$ and $\pi_0 = p_0$ is not an equilibrium. Furthermore, because $\eta(\pi_0, p_0) < 1$ implies $\eta(\pi_1, p_1) > 1$, it follows that $\tilde{\pi}_1(p_1, \sigma) > p_1$. Because $\tilde{\pi}_1$ is bounded from above, there exists a population equilibrium with a 'Non-Neutral Regime' in which $\pi_0 < p_0$. ■