# Social Status and Motivated Beliefs*

Víctor González-Jiménez†

September 14, 2021

**Abstract**

This paper shows that social status determines economic achievement by means of a psychological mechanism. Specifically, social status influences the way individuals form beliefs about their abilities and these beliefs are crucial for achievement. A theoretical framework formalizes the proposed mechanism and generates a set of testable predictions. Data from a cohort study and from two controlled experiments corroborate the validity of the theoretical predictions and, thus, of the proposed mechanism. This study highlights the role of social status in creating constraints that are internal to the individual and that have the potential to impair economic success.

**JEL Classification** : C91, D91, D83, Z13

**Keywords:** Social Status, Beliefs, Economic Mobility, Economic Achievement, Experiments.

---

*I am indebted to Patricio S. Dalton for guidance and useful comments on this paper. I thank Charles Noussair, Wieland Müller, Ilyana Kuziemko, Jan Stoop, Florian Schuett, Gabriele Camera, Robert Dur, Arthur Schram, Georg Kirchsteiger, Phillip Külpmann, Cole Williams, and Joaquin Gomez-Minambres for helpful suggestions. I also thank participants at seminars at Tilburg University, Chapman University, the University of Arizona, the Swedish Institute for Social Research at Stockholm University, the University of Vienna, and Trinity College Dublin for their feedback. The experiments reported in this study were funded by CentERLab at Tilburg University

†Department of Economics, The University of Vienna. Oskar-Morgernstern-Platz 1. 1090, Vienna Austria (e-mail: victor.gonzalez@univie.ac.at).

# 1. Introduction

How does social status affect economic success? Economists have established that pecuniary aspects inherent to social status in early stages of life are important for explaining economic outcomes at adulthood.[1] This traditional approach to understanding the long-lasting consequences of social status focuses on the role of *external* constraints. That is, it underscores how having limited access to markets of goods and services impairs economic success. I propose a novel mechanism in which social status confers advantages and disadvantages through constraints that are *internal* to individuals.[2] Specifically, social status, above and beyond the material advantages that it entails, induces beliefs that can constrain or encourage economic achievement.

Central to my proposal is the notion that performance on productive tasks and the beliefs that individuals hold about their ability to perform these tasks are positively related (See Benabou and Tirole (2002), Compte and Postlewaite (2004), Koszegi (2006), and Dalton et al. (2016) for evidence on this complementarity). The achievement of economic outcomes is thus not only determined by abilities but can be reinforced or undermined by beliefs. Sophisticated individuals internalize this relationship and, when advantageous, use social status to form and maintain potentially incorrect but favorable beliefs. For example, high beliefs can be held by low ability individuals when they attribute their high social standing to their abilities aided by prevalent conceptions about meritocracy in the society, i.e. the extent to which status reflects ability.[3] That these profitable self-serving beliefs can be formed introduces a disadvantage for low status individuals, who cannot form high beliefs in a similar vein and achieve lower economic outcomes. The existence of this mechanism has powerful implications. For instance, members of groups with long historical records of discrimination in a society have no other option but hold beliefs that make their disadvantage self-fulfilling.

Throughout this paper I adhere to Ridgeway and Walker (1995)'s definition, who characterize social status as *"the rank between individuals associated with prestige"*. According to this definition, social status refers to a person's relative standing in the society that involves social recognition, but does not necessarily entail pecuniary advantages and is not necessarily earned. Traditional economic theory predicts that these *mere* social ranks do not influence the economic achievement since they do not necessarily encompass external constraints.

[1] For example, parents' income and education (Sacerdote, 2002, Plug and Vijverberg, 2003, 2005), pre-natal and neo-natal care and conditions (Field et al., 2009, Bharadwaj et al., 2013), and neighborhood during childhood (Borjas, 1995, Chetty et al., 2015, Chetty and Hendren, 2018).

[2] An extended definition and an informative discussions of internal and external constraints can be found in Ghatak (2015).

[3] Another way in which these incorrect beliefs can be formed consists on internalizing stories of successful individuals or role models that shared similar social backgrounds.

However, the results presented in this paper go against this prediction. I show theoretically and empirically that social status, even when they do not yield material advantages, influence economic achievement by means of the different beliefs that they engender.

To establish the existence of the proposed mechanism, I examine whether individuals with similar abilities differ in their beliefs and economic achievement when belonging to different social status groups. Hence, the frameworks considered in this paper take social status as an exogenous assignment and guarantee that external constraints do not influence the achievement of economic outcomes. Confining myself to settings with these properties allows me to provide simple theoretical conditions guaranteeing the existence of the proposed mechanism, as well as the possibility to easily test the conjectures of the model with empirical evidence. Nevertheless, these advantages come at the cost of generalizability and the present study should be regarded as a proof of concept of the suggested influence of social status.[4]

A theoretical framework presented in Section 2 formalizes the proposed mechanism. The theoretical model demonstrates how social status influence beliefs about ability despite individuals having received an informative signal that fully disclosed their ability. This result emerges because individuals deliberately incorporate a high status assignment as a signal of ability in their belief system if such inaccuracy is profitable. Importantly, these beliefs manipulations are not unconstrained in the model but result as an equilibrium outcome from an interaction within the individual. Having the chance to form these motivated beliefs leads to higher expectations about ability, which in turn boosts performance. Receiving a low status implies that these beneficial beliefs cannot be accessed, a disadvantage generated by a seemingly innocuous initial status disparity.

Section 3 presents an empirical analysis that is consistent with this mechanism and that externally validates it. I use data from the 1970 British cohort study, which follows all individuals born in Britain in a week of 1970 over most of their lifespan. I find that that the educational aspirations of participants at an adolescent age, a proxy for their performance beliefs, are strongly influenced by their social status at birth. A relationship that has long-lasting economic consequences as participants who were born in higher social status households exhibit higher aspirations, which translate into a higher chance of attaining a high status as adults.

To complement the survey analysis I run two laboratory experiments. Their design is described in Section 4. In both experiments participants were assigned to either a high status

[4]A more complete modeling framework should determine whether the proposed mechanism emerges in a setting where individuals have access to relevant markets of goods and services and in which status is endogenous. The most prominent difficulties of using those setups are that it is theoretically difficult to determine which markets to include in the model and it is empirically involving to distinguish the influence of internal or external constraints on economic achievement in such context.

treatment or a low status treatment. Belonging to the high status treatment entailed receiving recognition from other participants and a positional good, properties that are consistent with the adopted definition of social status. Moreover, a cognitively challenging task was implemented before and after the treatment assignment. The first implementation of the task served to classify participants according to their ability. The second implementation measured performance after social status is assigned and in a setup in which the accurate completion of the task is rewarded with monetary incentives. The participants' beliefs about how well they think they will do on the task were elicited on multiple occasions in a session. Finally, to enhance learning about ability on the task, participants were given frequent feedback about performance as well as the opportunity to access a ranking that informed them about how well they performed the task relative to other participants.

In the first experiment, subjects were *randomly* assigned high status. The main result of that experiment, presented in Section 5, is that low ability participants assigned to the low status treatment display low belief and performance levels, while participants of similar ability who were assigned to high status display high beliefs and performance levels. This result cleanly corroborates the proposed mechanism: individuals endowed with low status cannot reach high outcomes since they do not have the opportunity to form and maintain high beliefs about their ability. In the second experiment the assignment to high status was *meritocratic*, that is only participants who exhibited high performance in the first implementation of the task were given the high status. Analyses of the data show that participants of low ability display low belief and performance levels. A result that demonstrates that the differences found in the first experiment are driven by the status assignment.

Finally, data on the participants' decision to access the relative performance ranking conclusively show that social status, when favorable, is deliberately used by participants to form and maintain motivated beliefs. I find that those who benefited the most from the high status assignment are less likely to access this ranking. They boosted their motivation by deliberately misinterpreting the social status assignment and avoided acquiring information that could contradict those self-serving beliefs.

**Contribution to existing literature**

This paper contributes to multiple strands of literature. First, it contributes to the literature on confidence maintenance and motivated beliefs (Benabou and Tirole, 2002, Compte and Postlewaite, 2004, Mobius et al., 2014, Benabou, 2015). While the theoretical framework is an adaptation of Benabou (2015) and Benabou and Tirole (2002), its results add novel interpretations and attributes to these models. For instance, I show that another mechanism to induce motivated beliefs, next to imperfect recall and attention manipulation, is the deliberate misinterpretation of social constructs such as social status. Moreover, the experimental data present clean evidence of participants engaging in motivated beliefs. In particular, they suppress relevant information about their ability, e.g. frequent feedback on the task, and instead behaved as if social status was a signal of their ability when such assignment was favorable. These empirical findings support the results of Eil and Rao (2011) and Mobius et al. (2014) in a context in which social status is (mis)interpreted as news about ability.

Second, it contributes to the literature that studies the effect of social status on economic decision making in experimental settings. The experimental protocol presented in this paper is based on the design of Ball et al. (2001) and Ball and Eckel (1998), who developed an artificial status manipulation to investigate the effect of status on behavior in double action markets and the ultimatum game. Other studies have adopted this manipulation to study the effect of status on coordination games (Eckel and Wilson, 2007), charitable giving (Kumru and Vesterlund, 2010), and public good games (Eckel et al., 2010). A common finding in these papers is that status changes behavior in a way that leads high status individuals to obtain higher earnings. This paper is, to the best of my knowledge, the first to use Ball et al. (2001)'s status manipulation to investigate the effect of status on individual performance on effort tasks and beliefs.[5] As in previous studies, I find that high status individuals derive higher earnings since they attain higher performance on the task. Moreover, the data on beliefs allow me to corroborate the relevant conjecture posed by Ball et al. (2001) and Ball and Eckel (1998) that the status assignment changes the participants' expectations about reasonable economic outcomes.

Finally, this paper contributes to the recent literature in economics that studies the influence of psychological factors such as aspirations, beliefs, and self-esteem on economic achievement (Bowles et al., 2001, Blanden et al., 2007, Bernard et al., 2014, Bogliacino and Ortoleva, 2015, Dalton et al., 2016, Genicot and Ray, 2017, Janzen et al., 2017). I develop a novel theoretical model based on the theory motivated beliefs that can explain why individuals who initially belong to a low social status exhibit low economic achievement as a consequence of having low beliefs about their own capabilities. Besides some differences regarding the assumed degree of sophistication of individuals (Dalton et al., 2016) and preference representation (Genicot and Ray, 2017), the main difference with respect to previous theoretical papers is that the theory is accompanied by distinct empirical analyses that validate it.

[5] This property differentiates this paper from the experimental literature investigating the effect of *competition* for status on behavior (Hopkins and Kornienko, 2006, Kuhnen and Tymula, 2012, Charness et al., 2014). In my experiment, the status assignment is unanticipated by subjects and is not reassigned. There is thus no competition for status.

## 2. Theoretical framework

This section formalizes the proposed mechanism. The model is based on the theory of motivated beliefs (Benabou and Tirole, 2002, Benabou, 2015) and shares similarities with models of Bayesian persuasion (Kamenica and Gentzkow, 2011). To the best of my knowledge, this is the first model applying these frameworks to the context of economic achievement and social status. Its main innovation is to demonstrate that a high social position gives access to advantageous and encouraging self-serving beliefs. As a consequence, social position has the power to self-sustain through a psychological mechanism.

### The benchmark

Consider a risk-neutral individual facing a time horizon of three periods $t = 0, 1, 2$. The individual's problem consists of exerting effort on a productive task at $t = 1$ given some set of information received at $t = 0$, to enjoy the monetary benefits derived from production on the task at $t = 2$.

In particular, at $t = 0$ the individual receives accurate information about his ability. Receiving this information can be interpreted as obtaining reliable feedback about his past performance on the task. For simplicity, I consider two ability levels on the task: high and low. Formally, let ability be the variable $\theta_i \in \{\theta_H, \theta_L\}$ where $\theta_H > \theta_L$. Moreover, the distribution of ability is known: if an individual were to be drawn at random from the population he would be of high ability with probability $q \in [0, 1]$.

At $t = 0$ the individual also receives a status class that endows him with a social standing. There are two status classes in the society: high and low. I represent social status with the variable $\sigma_j \in \{\sigma_H, \sigma_L\}$. It is commonly known that if an individual were drawn at random from the population he would belong to the high status with probability $r \in (0, 1)$.[6] Throughout it is assumed $q \geq r$, high status is scarce.[7]

I assume that ability and status are correlated in a non-negative way. This assumption captures the idea that societies differ in the extent to which social position reflects ability. When this correlation is zero, ability is not relevant to obtain a high standing in the society. Instead, when this correlation is large, the society assigns initial status in a way that overly

[6] The fact that $r \in (0, 1)$ entails that not all individuals in the society can belong to the high status, maintaining the desirability of high social status.

[7] This way of modeling status is adequate in settings whereby status is assigned or exogenously given, but it is not in settings whereby status can be earned as it is the case in contests for status. The work by Auriol and Renault (2008) and Moldovanu et al. (2007) provide complete theoretical frameworks for contests for social status.

reflects ability. Formally, let

$$\begin{aligned}\text{prob}(\theta_H, \sigma_H) &= \text{prob}(\theta_H)\text{prob}(\sigma_H) + \text{corr}(\theta_H, \sigma_H) \\ &= qr + \epsilon,\end{aligned} \tag{1}$$

where the parameter $\epsilon \in [0, \bar{\epsilon}(q, r)]$ captures the correlation between having a high social position and having high ability.[8] The degree of correlation between low status and low ability is also captured by $\epsilon$. Hence, the correlation between social status and ability is symmetric.[9] I assume that the individual knows the precise magnitude of $\epsilon$.

After receiving that set of information, the individual decides at $t = 1$ on the amount of effort to exert in the productive task. Let effort be a binary variable $e \in \{e_H, e_L\}$, with $e_H > e_L \geq 0$. For simplicity, it is assumed throughout that $e_H - e_L = 1$.

The agent faces the following trade-off. On one hand, choosing high effort, $e_H$, generates disutility. I represent the cost of exerting high effort by the function $c(e)$, which is assumed to have the following functional form:

**Assumption 1.** $c(e) = \begin{cases} c & \text{if } e_H, \\ 0 & \text{if } e_L. \end{cases}$ *Where* $c > 0$.

On the other hand, exerting high effort implies receiving larger monetary rewards at $t = 2$, as it is assumed that effort raises output in a deterministic way and that output on this task is valuable. The following assumption presents the production function used in the model:

**Assumption 2.** $f(e, \theta_i) = \theta_i e$.

Finally, I assume that the individual can experience psychological utility. This component of utility captures the idea that maintaining high confidence about ability on the task generates utility gains when working on it (Compte and Postlewaite, 2004, Koszegi, 2006, Benabou, 2015). I model psychological utility by incorporating the individual's belief about his own performance in the utility function:

**Assumption 3.** $\psi(e, \theta_i) = \mathbb{E}_1(\theta_i e)$.

All in all, the utility of the individual at $t = 1$ can be written as:

$$U_1(e) = af(e, \theta_i) + s\psi(e, \theta_i) - c(e), \tag{2}$$

[8]The upper bound $\bar{\epsilon}(q, r)$ is defined by standard probability rules, that is $0 \leq prob(\theta_H, \sigma_H) \leq 1$. Hence, $\bar{\epsilon}(q, r) = (1 - q)r$.

[9]Symmetry emerges as a necessary condition required to respect the marginal probabilities $r$ and $q$ under the representation $prob(\theta_H, \sigma_H) = prob(\theta_H)prob(\sigma_H) + corr(\theta_H, \sigma_H) = qr + \epsilon$. Note that under symmetry $prob(\theta_H, \sigma_H) + prob(\theta_H, \sigma_L) = q = prob(\theta_H)$ and also that $prob(\theta_L, \sigma_H) + prob(\theta_L, \sigma_L) = 1 - q = prob(\theta_L)$.

where $a > 0$ represents a monetary amount symbolizing the rewards received in exchange of producing a level of output $y$, and $s \geq 0$ weights the impact of psychological utility on the agent's utility. Since the individual is fully informed about his ability, then psychological utility reduces to $\mathbb{E}_1(\theta_i e) = \theta_i e$, and equation (2) becomes:

$$U_1(e) = (a + s)\theta_i e - c(e). \tag{3}$$

From equation (3) it can be established that the individual chooses $e_H$ if such choice is profitable. That is whenever $U_1(e_H) \geq U_1(e_L) \Leftrightarrow (a + s)\theta_i \geq c$. Instead, if the cost of effort is such that $c > (a + s)\theta_i$, low effort is chosen. In the remainder of the model, I focus on the interval $c \in \Big((a + s)\theta_L, (a + s)\theta_H\Big]$. This interval encompasses the relevant case in which, absent self-deception, only high ability individuals choose high effort.

Note that the model, as it is, predicts that the individual's decision to exert high effort is not affected by his social status. That is because $\sigma_j$ has no economic or psychological value and does not yield any advantage to perform the task. Therefore, in a framework in which individuals achieve outcomes by means of their own ability and effort, an initial status assignment does not affect economic outcomes.

### Motivated beliefs from self-deception

I depart from the benchmark by introducing self-deception in the model. That is, I let the individual deliberately hold erroneous beliefs about his own ability. However, instead of letting the individual freely manipulate his beliefs, I model self-deception as a strategic interaction within the individual. As a consequence, incorrect beliefs about ability must emerge as an equilibrium outcome.

Assume that there are two selves within the individual. *Self 0*, an informed self who knows his ability, $\theta_i$, and social status, $\sigma_j$, and *Self 1* an uninformed self who has the task of choosing $e$ without knowing $\theta_i$ nor $\sigma_j$. The interaction between these two selves consists on the transmission of relevant information: Self 0 uses available information to send an ability signal to Self 1. Denote this signal by $\tilde{\theta}_k \in \{\tilde{\theta}_H, \tilde{\theta}_L\}$, which may be truthful, $k = i$, or may be distorted, $k = j$ with $j \neq i$. Distortions of the truth or lying can be interpreted as Self 0 forgetting unfavorable signals, misinterpreting favorable signals, or paying too much attention to favorable signals while paying too little attention to unfavorable signals.[10] The assumed information asymmetry between the selves generates a strategic interaction. To derive higher monetary rewards or generate higher psychological utility from holding high

[10]In the words of Benabou (2015) "We thus allow the agent to process good and bad signals asymmetrically in term of attention, interpretation, memory or awareness." This model also allows for these traditional channels of motivated beliefs.

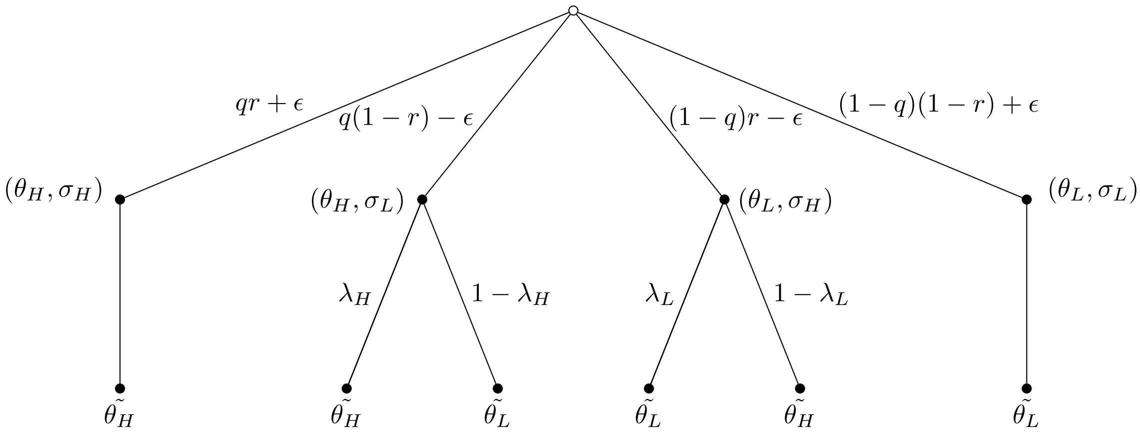


Figure 1: Self 1's reaction to signals

beliefs, Self 0 can send signals that induce inaccurate beliefs and affect Self 1's choice of effort. Self 1 anticipates these intentions and discounts the received signals.[11]

The strategy of Self 0 consists on telling the truth with probability $\lambda_k \in [0,1]$ with $k = \{i, j\}$, so mixed strategies of truth-telling are allowed. Importantly, and as it is standard in the literature of motivated beliefs, I assume that lies cannot be *ex nihilo* fabricated, and truth distortions need to be supported by a social status assignment. Hence, the strategy space of Self 0 is the range of the function $\mathcal{S}$, defined as $\mathcal{S} : \{\theta_L, \theta_H\} \times \{\sigma_L, \sigma_H\} \to \{1, \lambda_L, \lambda_H, 1\}$, where $\lambda_L$ and $\lambda_H$ are the truth-telling probabilities of individuals with ability-status tuples $(\theta_L, \sigma_H)$ and $(\theta_H, \sigma_L)$, respectively.

*Self 1* makes inference about the veracity of $\tilde{\theta}_k$ using Bayes' rule. When $\tilde{\theta}_k$ is received, he acknowledges that with probability $\lambda_k \in [0,1]$ Self 0 is telling the truth and he is of ability $\theta_k$ and that with probability $1 - \lambda_k$ Self 0 is lying, and he is instead of ability $\theta_{-k}$, the ability level corresponding to the signal *that was not sent* by Self 0. Figure 1 illustrates Self 1's updating process.

In particular, Self 1's posterior probability that a favorable signal $\tilde{\theta}_H$ reflects his ability is:

$$p_H(\lambda_H, \lambda_L) = \frac{qr + \lambda_H q(1-r) + \epsilon(1-\lambda_H)}{qr + \lambda_H q(1-r) + \epsilon(1-\lambda_H) + \chi\left(r(1-q) - \epsilon\right)(1-\lambda_L)}. \quad (4)$$

[11] To build intuition about the incentives that Self 0 faces, consider an scenario in which an agent received at $t = 0$ a low ability and high status tuple: $(\theta_L, \sigma_H)$. In such case, Self 0 could either send $\tilde{\theta}_L$, a signal reflecting his true ability, or $\tilde{\theta}_H$, a distorted signal. Note that the psychological component in the individual's utility, $\mathbb{E}_1(\theta_i e)$, becomes higher when Self 1 believes that his ability are high. However, these beliefs also lead to steeper costs of effort if those higher beliefs lead Self 1 to choose $e_H$. Thus, depending on how costly is to choose the high level of effort, Self 0 could be better off distorting information.

Similarly, Self 1's posterior probability that an unfavorable signal $\tilde{\theta}_L$ reflects his ability is:

$$p_L(\lambda_H, \lambda_L) = \frac{(1-q)(1-r) + \lambda_L(1-q)r + \epsilon(1-\lambda_L)}{(1-q)(1-r) + \lambda_L(1-q)r + \epsilon(1-\lambda_L) + \chi\left(q(1-r) - \epsilon\right)(1-\lambda_H)}. \tag{5}$$

These equations feature an important intuition of the model. Equation (4) shows that the posterior $p_H(\lambda_H, \lambda_L)$ crucially depends on $\lambda_L$. Specifically, if a Self 0 with low ability *never* engages in self-deception, so that $\lambda_L = 1$, a favorable signal, $\tilde{\theta}_H$ is credible and is regarded as truthful. Hence, $p_H(\lambda_H, 1) = 1$. In contrast, if a Self 0 with low ability *always* engages in self-deception, a favorable signal is not credible and Self 1 assesses the probability of having high ability according to his prior.[12] Similarly, equation (5) shows that the posterior $p_L(\lambda_H, \lambda_L)$ decreases with higher values of $\lambda_H$.

We are now in a position to describe the way in which Self 1 forms beliefs upon receiving a signal $\tilde{\theta}_k$. Equations (4) and (5) imply that Self 1's beliefs about his ability are given by the following expectation,

$$\mathbb{E}_1\Big(\theta_k|\tilde{\theta}_k\Big) = p_k(\lambda_H, \lambda_L)\theta_k + \Big(1 - p_k(\lambda_H, \lambda_L)\Big)\theta_{-k}. \tag{6}$$

Given these beliefs, Self 1's program is:

$$\max_{e \in \{e_H, e_L\}} \quad (a+s)\mathbb{E}_1\Big(\theta_k|\tilde{\theta}_k\Big)e - c(e) \tag{7}$$

Lemma 1 presented in Appendix A provides the solution to Self 1's program. It proves the existence of a threshold probability $\hat{\lambda}_{-k}$ that makes Self 1 indifferent between choosing $e_H$ or $e_L$. This lemma builds upon the aforementioned dependence of posterior probabilities on the magnitudes $\lambda_H$ and $\lambda_L$. Specifically, high effort is chosen when a favorable signal is received and low ability individuals tell the truth at a rate higher than the threshold $\hat{\lambda}_H$, or alternatively when an unfavorable signal is received and high ability individuals tell the truth at a rate lower than the threshold $\hat{\lambda}_L$. Favorable signals need to remain credible or unfavorable signals are know to be bogus.

Before analyzing the optimal choice of Self 0, I impose the assumption that this self faces a cost $m(\lambda)$ when engaging in self-deception. Intuitively, forgetting, misinterpreting, and deliberately paying too little or too much attention to the received information is cognitively costly. For simplicity, the following cost schedule is assumed.

**Assumption 4.** $m(\lambda) = \begin{cases} m \ \textit{if} \ \lambda_i < 1, \\ 0 \ \textit{if} \ \lambda_i = 1. \end{cases}$ *With* $m > 0$ *and* $\lambda_i = \{\lambda_L, \lambda_H\}$.

[12]The specific value of the posterior when $\lambda_L = 0$ is $p_H(\lambda_H, 0) = \frac{qr + \lambda_H q(1-r)) + \epsilon(1-\lambda_H)}{qr + \lambda_H q(1-r)) + \epsilon(1-\lambda_H) + \chi(r(1-q) - \epsilon)}$, which corresponds to his prior when the received signal is not informative.

All in all, Self 0's program can be written as:

$$\max_{\lambda_i \in [0,1]} \quad \mathbb{E}_0\Big(\mathbb{E}_1\left(U_1(e,\lambda_i)\right)\Big) - m(\lambda) \tag{8}$$

A comparison of Self 0 and Self 1 programs' formally demonstrates their strategic interaction. As mentioned before, this strategic interaction is caused by an information asymmetry, which can lead them to hold different beliefs. When receiving a favorable signal, Self 1's beliefs about his ability are equal to $\mathbb{E}_1(\theta_i) = \Big((\theta_H - \theta_L)p_H(\lambda_H, \lambda_L) + \theta_H\Big)$ and Self 0's beliefs about Self 1's beliefs are equal to $\mathbb{E}_0(\mathbb{E}_1(\theta_i)) = \lambda_i\theta_i + (1-\lambda_i)\Big((\theta_H - \theta_L)p_H(\lambda_H, \lambda_L) + \theta_H\Big)$. These two beliefs are different unless $\lambda_i = 0$. Since beliefs enter the utility function instrumentally, that the selves hold different beliefs implies that they maximize different objective functions.

The game between the two selves is solved using subgame perfect Nash equilibrium. The relevant equilibria of the game are defined in Definition 1 and Definition 2, and their existence is guaranteed by Proposition 1 and Proposition 2. The proofs of the main theoretical results are relegated to Appendix A. Other resulting equilibria of the game and their proofs are relegated to Appendix B. I start by defining and proving the existence of an equilibrium in which social status determines the effort choice and beliefs of low ability individuals.

**Definition 1.** *A semi-pooling equilibrium of the status and self-deception game is characterized by the tuple* $(\lambda_p^{**}, e_p^{**})$*, where:*

$$e_p^{**} = \begin{cases} e_L & \textit{if } (\theta_L, \sigma_L), \\ e_H & \textit{if } (\theta_L, \sigma_H) \textit{ or } \theta_H, \end{cases}$$

*and*

$$\lambda_p^{**} = \begin{cases} 1 & \textit{if } (\theta_L, \sigma_L) \textit{ and } \theta_H, \\ \hat{\lambda}_L & \textit{if } (\theta_L, \sigma_H). \end{cases}$$

**Proposition 1.** *Under Assumptions 1-4, a semi-pooling equilibrium is sustained if* $s > 0$*,* $c \leq \bar{c}$*, and* $m \leq \bar{m}$*, where* $\bar{c} := (a+s)\left(\frac{se_Hq\theta_H + \chi((1-q)r-\epsilon)(m-(a+s)\theta_L)}{se_Hq + (a+s)\chi((1-q)r-\epsilon)}\right)$ *and* $\bar{m} := \frac{se_Hq(\theta_H - \theta_L)}{q + \chi(r(1-q)-\epsilon)}$.

The equilibrium proved by Proposition 1 features a low ability individual who exerts high effort if he belongs to the high status, but exerts low effort otherwise. This equilibrium emerges and is sustained in the following way. To derive utility gains from psychological utility, Self 0 with tuple $(\theta_L, \sigma_H)$ engages in a strategy that maintains the credibility of favorable signals and as a consequence incentivizes $e_H$ when a favorable signal is sent. In particular, he sends favorable signals at the rate $1 - \hat{\lambda}_L$, which according to Lemma 1 makes

Self 1 indifferent between choosing high and low effort. This constitutes a profitable strategy for Self 0 since the cost of lying is not too high ($m < \bar{m}$) and psychological utility is kept at maximum. Self 1 responds to receiving a favorable signal by choosing high effort, a profitable strategy since the cost of effort is not too high ($c < \bar{c}$) and psychological utility is kept at maximum. Given this behavior of low ability individuals with high status, high ability individuals understand that it is best-strategy is to be truthful inasmuch as favorable signals are credible and self-deception is costly, $m > 0$.

It is of special interest to understand how the magnitude of $\epsilon$, the correlation between status and ability, affects the emergence of the semi-pooling equilibrium proved by Proposition 1. To that end, I first establish the conditions required for this equilibrium to emerge under the special case $\epsilon = 0$.

**Corollary 1.** *Under $\epsilon = 0$, the equilibrium in Proposition 1 emerges if $(a+s)\theta_L < m < \bar{m}\big|_{\epsilon=0}$.*

When status and ability are independent, receiving a high status carries no information about ability. Therefore, forming motivated beliefs by misinterpreting status as an indicator of ability is not possible. However, having received the high status enables the low ability individual to engage in motivated beliefs through standard channels: forgetting, paying little attention and/or misinterpreting the unfavorable signal that fully concealed his type, $\theta_L$. That these standard channels of truth distortion are available is a direct consequence of the assumed restrictions on Self 0's strategy space, i.e. that lying needs to be supported by a status signal. Hence, the high status assignment enables this individual to form inaccurate beliefs if they are profitable.

Next, I show that the semi-polling equilibrium in Corollary 1 emerges under the most stringent conditions. This is because larger values of $\epsilon$ considerably facilitate the formation of profitable motivated beliefs.

**Corollary 2.** *Higher values of $\epsilon$ increases the value of thresholds $\overline{c}$ and $\overline{m}$. The equilibrium in Proposition 1 emerges under less stringent conditions the higher $\epsilon$ is.*

The intuition behind this comparative static is simple: larger values of $\epsilon$ imply that individuals more often receive a social status that matches their ability. As a consequence, individuals who did not can more easily convince themselves that the assigned status reflects their ability, this makes the self-deception process less costly. The traditional channels of motivated beliefs are reinforced by the ease at which status can be misinterpreted as an indicator of ability.

An alternative interpretation of Corollary 2 relaxes the assumption that $\epsilon$ is exactly known. In this setting, the individual can more easily form high beliefs after receiving the high status when the value of $\epsilon$ is overestimated. This interpretation is appealing as it implies that self-serving beliefs are more likely to emerge when the correlation between ability and status is overestimated in the society due to conventions about the extent to which social position reflects ability.

Alternatively, when the cost associated to the self-deception strategy is high, individuals are better off self-signaling their true ability. Proposition 2 presents the conditions for the emergence of this *separating equilibrium*.

**Definition 2.** *A separating equilibrium of the status and self-deception game is characterized by the tuple* $(\lambda_s^{**}, e_s^{**})$*, where*

$$e_s^{**} = \begin{cases} e_L & \text{if } \theta_L, \\ e_H & \text{if } \theta_H, \end{cases}$$

*as well as*

$$\lambda_s^{**} = \lambda_L^{**} = \lambda_H^{**} = 1$$

**Proposition 2.** *Under Assumptions 1-4, there exists a separating equilibrium that is sustained if* $s > 0$ *and* $m > \bar{m}$*, where* $\bar{m} := \frac{s e_H q(\theta_H - \theta_L)}{q + \chi(r(1-q) - \epsilon)}$.

The equilibrium proved by Proposition 2 entails that social status has no economic consequences. That is because self-deception is not profitable. In that case, the ability parameter fully determines whether high or low effort is exerted.

To conclude, the model generates two competing results. Proposition 1 demonstrates that social status generates beliefs and performance differences among similarly skilled individuals as a result of motivated beliefs being profitable. This equilibrium is in line with my proposal. In contrast, Proposition 2 and the benchmark of the model show that social status, as defined in this paper, is not powerful enough to yield relevant economic consequences. The question of whether one or the other equilibrium governs behavior will be empirically investigated with an analysis of survey data presented in Section 3 as well as two controlled laboratory experiments described in Sections 4 and 5.

## 3. Survey evidence

This section presents survey evidence that is consistent with Proposition 1. I collect data from the British cohort study, which follows *all* individuals born in the United Kingdom in the third week of April 1970. That study is composed by different waves or follow-ups, I use the data gathered in the waves corresponding to the years 1970, 1975, 1980, 1986, and

2008. From these waves I obtain, among others, variables that capture the participants' social status at birth, their educational aspirations at an adolescent age—used as a proxy for beliefs about ability—, and their social status at an adult age. To test whether Proposition 1 holds, I investigate whether social status at birth significantly affects aspirations, and whether this relationship is powerful enough to influence the participants' social status at adulthood.

An advantage of this approach is that it allows me to evaluate whether the proposed mechanism also materializes outside of the laboratory and whether it is powerful enough to affect the life of individuals. A major difference with most studies on the topic is that they focus on short-term economic outcomes or proxies for economic achievement such as for future-oriented behavior variables (Macours and Vakis, 2014, Bernard et al., 2014, Laajaj, 2017, Janzen et al., 2017, Ross, 2019). To my knowledge, this is the first study that evaluates the existence and extent of this relationship in a developed country and for the longest documented time span, with more than 30 years separating the measurement of these variables. Finally, the data are rich and contain variables that enable to account for factors that have been previously proved to affect economic achievement.

Table 1 presents the description and descriptive statistics of the key variables used in the analyses. They are all ordinal and take values from one to five. *"Status Parents"* represents social status, i.e. $\sigma_j$ in the model. It captures the parents' type of occupation when participants are born, a social status representation often used in sociology and a good proxy for parents' income. This variable takes a value of five for professional occupations, four for managerial and technical occupations, three for non-manual and partly-skilled occupations without a degree, two for manual partly-skilled occupations without a degree, and one for unskilled occupations. *"Status Adult"* represents economic achievement, i.e. $\theta_i e$ in the model. It captures the participants' type of occupation at an adult age. The categories of this variable are constructed to exactly match that of Status Parents, so that comparisons between these two variables are feasible and coherent. Finally, *"Aspirations"* represents beliefs, $\mathbb{E}_1(\theta_i e)$. This variable captures the participant's educational aspirations at age 16, which in turn reflect their occupational aspirations. Aspirations are not exactly the same as beliefs, as they are not probabilistic assessments of outcomes but instead forward-looking goals that would only be possible if constraints were lifted or if individuals changed their behavior. However, aspirations are a good proxy for beliefs since, in the words of Bernard et al. (2014), *"the beliefs held by individuals about their environment and themselves influence their aspirations.* For the sake of comparability, the scale of Aspirations is constructed so that it exactly matches that of the other two status variables.[13]

[13]To make a closer link to the theoretical model, I also use discrete versions of these variables. The variables *"High Status Parents"*, *"High Aspirations"*, and *"High Status"*, indicate either occupations that require non-vocational tertiary education levels such as university, or the aspiration to achieve these education

Table 1: Descriptive statistics of variables

| **Variable** | Type | Scale | Mean | St. Dev. | Median |
|---|---|---|---|---|---|
| Aspirations | Ordinal | 1 to 5 | 3.386 | 1.283 | 3 |
| Status Parents | Ordinal | 1 to 5 | 2.931 | 0.873 | 3 |
| Status Adult | Ordinal | 1 to 5 | 3.323 | 0.831 | 3 |

Note: This table presents the type, scale, average, standard deviation, and median of the main variables used in the survey analysis. "Status Adult" is an ordinal variable that captures the individual's occupation at the age of 37. "Aspirations" is an ordinal variable that captures the educational aspirations of the individual at adolescence. "Status Parents" is an ordinal variable that captures the parents' occupation at birth. These variables all take a value of 5 for professional occupations, 4 for managerial and technical occupations, 3 for non-manual partly-skilled occupations without a degree, 2 for manual partly-skilled occupations without a degree, and 1 for unskilled occupations.

The empirical validity of Prediction 1 is examined using regressions that relate these three variables. To control for factors that have been previously found to influence economic achievement, these regressions include an extensive set of covariates. For instance, cognitive abilities (Cunha et al., 2010) are taken into account in the model with the participants' scores on cognitive, math, and reading tests scores during early and late infancy. Non-cognitive abilities (Heckman et al., 2006) with the participant's personality traits, self-esteem, and locus of control scores measured during adolescence (Bowles et al., 2001). These variables seek to capture the participant's ability, $\theta_i$ in the model.[14] I also control for internal constraints that have been previously established to be important for achievement but that are not part of the model, such as the parents' expectations and teachers' expectations about the participants' diligence in school at adolescence (Besley, 2016, Jensen, 2010). Furthermore, external constraints are captured by the inclusion of a variable that captures the participant's family income in 1980 and a variable that captures the type of neighborhood in which participants lived in their adolescence (Chetty and Hendren, 2018, Chetty et al., 2015). Appendix C provides the full description and descriptive statistics of all variables used in the analyses.

I first regress Aspirations on Status Parents while controlling for the aforementioned set of variables. The estimates of ordered logistic regressions are presented in columns 1,2, and 3 of Table 2. I find that the average individual in the sample aspires to the highest education level, a career that requires a university degree, with 28% chance. Moreover, being born in a

levels. In the interest of space, I relegate the analysis using the binary version of the relevant variables to Appendix D.1.

[14]A significant and correlation between these variables and Status Parents, implies $\epsilon > 0$. I report that this is indeed the case for all variables.

household belonging to the highest socio-economic status is associated with a 9.7% higher probability to aspire to the highest education level as compared to being born in a household belonging to the lowest socio-economic status. A difference that is significant at the 1% significance level. This result corroborates previous research showing that higher income and greater economic resources are positively correlated with higher aspirations (Chiapa et al., 2012, Macours and Vakis, 2014, Ross et al., 2021). The main difference with respect to those studies is that these data capture a long-term relationship, spanning over 16 years.

Next, I investigate whether this relationship between status and aspirations has consequences on status at an adult age, the variable that captures economic achievement. I regress the socio-economic status of individuals at the age of 37 on educational aspirations, socio-economic status at birth, the interaction between these two variables, and the set of control variables. Columns 4, 5, and 6 in Table 2 present the estimates of ordered logistic regressions. The coefficient corresponding to the interaction between Aspirations and Status Parents is statistically significant at the 5% level for all specifications. Higher social status at birth and higher aspirations are associated with higher social status at adulthood, ceteris paribus. This finding is consistent with Proposition 1, and with previous research in developing countries showing that aspirations and status jointly influence economic outcomes (Macours and Vakis, 2014, Janzen et al., 2017, Ross, 2019). Furthermore, the coefficients associated to Status Parents Aspirations suggest that higher aspirations and higher social status at birth, on their own, do not correlate with higher achieved status. Appendix D.1 shows that the estimates of regressions using the discrete versions of the relevant variables yield similar conclusions.

To illustrate the far-reaching consequences of the proposed mechanism, I quantify how changes in aspirations and social status affect economic achievement, ceteris paribus. That exercise is performed with the estimates presented in column 6 of Table 2. Having the highest aspiration level is 9.4% more likely to yield the highest status at adulthood when an individual with average characteristics is born in a household belonging to the highest status as compared to a household in the lowest status. Moreover, being born in the highest status is 12.77% more likely to yield the highest social status at adulthood when an individual with average characteristics sets the highest, rather than the lowest, aspiration level.

While these findings are consistent with the proposed mechanism, they cannot be treated as conclusive evidence of Proposition 1. There are two compelling reasons to consider these results suggestive rather than corroborative. First, this data analysis is silent about the precise psychological mechanism by which aspirations and initial social status relate, and is thus insufficient to demonstrate that social status is deliberately used by individuals to form motivated beliefs. Hence, these results can also be explained by other theoretical frameworks such as that of Dalton et al. (2016). This is a disadvantage of using observational data,

Table 2: Determinants of aspirations and achieved social status

| | (1) Aspirations | (2) Aspirations | (3) Aspirations | (4) Status Adult | (5) Status Adult | (6) Status Adult |
|---|---|---|---|---|---|---|
| Status Parents | 0.462*** | 0.200*** | 0.148*** | -0.043 | -0.197 | -0.229 |
| | (0.043) | (0.052) | (0.053) | (0.150) | (0.184) | (0.185) |
| Aspirations | | | | 0.204 | 0.080 | -0.012 |
| | | | | (0.133) | (0.162) | (0.163) |
| Aspirations*Status Parents | | | | 0.112*** | 0.103** | 0.103** |
| | | | | (0.040) | (0.048) | (0.049) |
| Mother's age at birth | 0.005 | -0.003 | -0.001 | 0.005 | 0.002 | 0.004 |
| | (0.004) | (0.005) | (0.004) | (0.005) | (0.005) | (0.006) |
| Female | 0.317*** | 0.211** | 0.122 | -0.499*** | -0.408*** | -0.470*** |
| | (0.076) | (0.090) | (0.100) | (0.087) | (0.108) | (0.121) |
| Family income 1980 | | -0.114*** | -0.096*** | | -0.077*** | -0.082*** |
| | | (0.025) | (0.025) | | (0.028) | (0.028) |
| Constant | 0.631** | 0.809 | -0.160 | -2.599*** | -1.991* | -1.773 |
| | (0.272) | (0.660) | (0.768) | (0.583) | (1.040) | (1.160) |
| Cognitive skills | NO | YES | YES | NO | YES | YES |
| Non-cognitive traits | NO | YES | YES | NO | YES | YES |
| Self-reported skills | NO | NO | YES | NO | NO | YES |
| Neighborhood | NO | NO | YES | NO | NO | YES |
| Expectations | NO | NO | YES | NO | NO | YES |
| Attitudes toward school | NO | NO | YES | NO | NO | YES |
| Health | NO | NO | YES | NO | NO | YES |
| Country Fixed effects | YES | YES | YES | YES | YES | YES |
| Region Fixed effects | YES | YES | YES | YES | YES | YES |
| Log-Likelihood | -3640.820 | -2770.929 | -2649.116 | -2338.892 | -1812.992 | -1769.918 |
| N | 2661 | 2178 | 2178 | 2,090 | 1,719 | 1719 |

Note: Columns 1, 2, and 3 presents the estimates of the ordered logistic regression of the model $Aspirations_i = \beta_0 + \beta_1 StatusParents_i + Controls'\Gamma + \epsilon_i$ with $\epsilon_i \sim logistic$. Columns 4, 5, and 6 presents the estimates of the ordered logistic regression of the model $StatusAdult_i = \beta_0 + \beta_1 StatusParents * Aspirations + \beta_2 Aspirations + \beta_3 StatusParents_i + Controls'\Gamma + \epsilon_i$ with $\epsilon_i \sim logistic$. "Status Adult" is an ordinal variable that represents the individual's occupation at the age of 37. "Aspirations" is an ordinal variable that captures the educational aspirations of the individual at adolescence. "Status Parents" is an ordinal variable that represents the parents' occupation at birth. The description of the control variables is presented in Appendix C. Clustered standard errors at the individual level in parentheses. *** denotes significance at the 0.01 level, ** denotes significance at the 0.05 level, * denotes significance at the 0.1 level.

which typically allow the researcher to establish externally-valid relationships but are often insufficient to distinguish and determine the validity of competing psychological mechanisms.

Second, stringent econometric assumptions are required in order to i) interpret the obtained estimates in a way that the predictions of Proposition 1 are validated, and ii) guarantee that these estimates are unbiased. Regarding i), the estimates presented above can be interpreted as marginal effects under the assumption of *predeterminedness*, that is that social status at birth affects aspirations but not the other way around. A simple intergenerational bequest argument can dispute this assumption, namely that parents achieve a high social status to endow children with the possibility of having high aspirations. On ii), there might be unobserved factors that affect achievement and aspirations and that are not being taken into account in the analysis, yielding biased estimates due to omitted variable bias.

To complement these findings, the remainder of the paper is devoted to investigate the validity of Proposition 1 in a controlled laboratory setting. A set of laboratory experiments allow me to overcome the problem of omitted variable bias inasmuch as social status in the laboratory can be randomly assigned. Moreover, that this randomization is feasible allows me to interpret changes in clean measures of performance and beliefs as the direct consequence of the social status assignment. Finally, data on the subjects' willingness to demand valuable information about ability allow me to establish whether social status is used by individuals to hold motivated beliefs. Section 4 presents the design of the experiments and Section 5 presents the results.

## 4. Experimental design and procedures

### 4.1. Experiment 1

The experiment was conducted at Tilburg University's CentERLAB. Participants were students at the university and were recruited through an online system. The data consist of 8 sessions with a total of 136 participants. On average a session lasted approximately 60 minutes. Between 13 and 24 participants took part in a session. The currency used in the experiment was euros. I used Z-Tree (Fischbacher, 2007) to implement and run the experiment. Participants earned on average 11.55 euros. The instructions of the experiment are presented in Appendix F.

Figure 2 provides a timeline of the experiment. That figure facilitates the description of the design and presents, in its lower part, how each part of the experiment relates to the model. As shown in the figure, the experiment consisted of two parts: Part 1 and Part 2. Upon arrival participants were explicitly informed about this composition of the experiment.

In both parts, participants were invited to complete different versions of the Raven's Matrices test, a task that demanded cognitive resources (Raven, 1989). Hence, it was more difficult for motivated but unskilled participants to improve on their performance as compared to a task that, say, only demanded attention or effort. Furthermore, this task was chosen to constitute an stringent test of Proposition 1. Exerting high effort entailed a high cost, $c$, and, since the task demanded cognitive resources from subjects, it made the cost of distorting information, $m$, steeper.

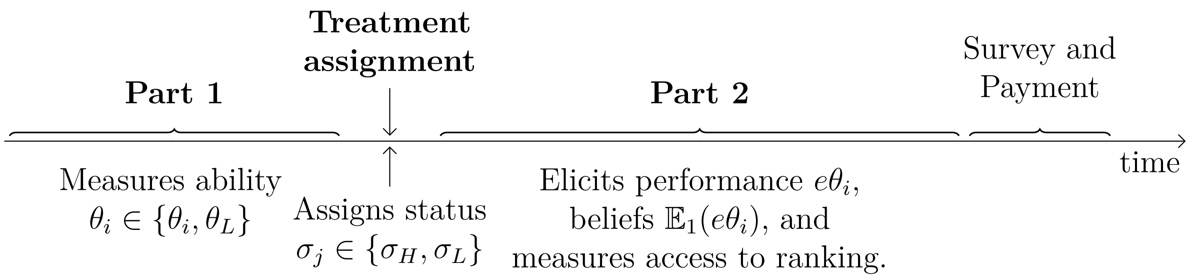


Figure 2: Timeline of the experimental design

In Part 1 of the experiment, Set I of the Advanced Progressive Matrices test (APM) was implemented. This "quick version" of the Raven's test consists of 12 matrices that cover the full range of difficulty in the Raven's test. In fact, performance in this set has been proven to be a predictor of performance in the complete version of the test Bors and Stokes (1998). Participants had five minutes to complete this number of matrices, as recommended by Raven (1989). Although participants did not face monetary incentives in this part of the experiment, they were encouraged to do their best. The aim of this part of the experiment was to measure their ability. As shown in the lower part of Figure 2, Part 1 is used to determine if a participant had high $\theta_H$ or low ability $\theta_L$ on the task. This classification will be described in more detail in the next section. Once the specified time for this part of the experiment was over, feedback about performance, i.e. the number of correctly solved matrices, was given in private to each participant.

After Part 1 of the experiment was completed, I introduced status differences between subjects participating in the same session. The well-known, clear, and simple status manipulation developed by Ball and Eckel (1998) and Ball et al. (2001) was implemented. Specifically, participants were *randomly* assigned to one of two treatments: the *"High Status"* treatment, in which they received a symbolic award and social recognition from their peers, or the *"Low Status"* treatment, where they did not receive the award nor social recognition.[15]

[15]The rationale behind inducing an artificial status rank rather than using a naturally occurring status rank was to minimize the possibility of disagreements that may arise from using the latter type of status. For instance, subjects may disagree about the rank of an status allocation based on academic performance

These treatments correspond to $\sigma_H$ and $\sigma_L$ in the model as shown by Figure 2.

The random assignment to the treatments enables performance and belief comparisons between participants with similar ability but assigned to different treatments. Although participants were not informed about this exact treatment assignment rule, they were also not explicitly lead to believe that the assignment had to do with their performance. The exact wording of the message given and read to participants was: *"The following participants were assigned to the GOLD group. Please come up to the front as we call your ID number and receive your medal"*. The experimenter reports that no participant questioned the reason for the assignment to the treatments. According to the model, the random treatment assignment implies $\epsilon = 0$. However, that subjects are not explicitly informed about this rule might facilitates the formation of motivated beliefs since they can assume $\epsilon > 0$ (See Corollary 2 and the intuition given after that corollary).[16]

In Part 2 of the experiment, I implemented Set II of the APM. This set is the complete version of the Raven's test and consisted of 36 matrices. Additionally, the most difficult 24 matrices of the Standard Progressive test (SPM) were included.[17] In total, participants had 20 minutes to solve as many matrices as they could and faced a monetary incentive of 0.5 euros for each correctly solved matrix. This part of the experiment was introduced to measure performance, $f(e, \theta) = \theta_i e$ in the model, as shown in Figure 2. Note that by explicitly telling subjects that the experiment consisted of two parts and by giving full instructions at the beginning of each part, I eradicate the possibility of subjects expecting status to be reassigned on the basis of their performance on the task in Part 2.

The pre-specified time given to participants in Part 2, also recommended by Raven (1989), was divided in five rounds of four minutes each. This division of the total time into rounds had multiple purposes. First, it allowed me to provide individual feedback on the task at the end of each round, i.e. the number of correctly solved matrices in previous rounds. Providing

(what kind of courses are considered for this rank?), gender (which characteristics make one gender rank higher than the other?), or socio-economic status (I am right now having similar status than someone else of my age that is working, but I am investing in education at the moment). See Ball et al. (2001) for a more comprehensive discussion about the rationale of this status differential, as well as a more detailed discussion of its implementation.

[16]I favor this design instead of one in which subjects are explicitly told that the treatment assignment is random. First, because this is seldom the case: one's social position can be determined by luck or one's ability and it is hard to disentangle the exact contribution to each of these factors, which provides room for interpretation and self-serving belief formation. Second, because this experiment seeks to provide a first proof of concept that social status *can* trigger internal constraints. Not making the random assignment explicit eases the emergence of motivated beliefs because it gives subjects some leeway to interpret their received social status. However, note that this design still requires subject to engage in a considerable degree of self-deception, since they need required to suppress or avoid relevant feedback in order to form and maintain the desired beliefs.

[17]The SPM is also a Raven's matrices test but with a lower average difficulty as compared to the APM. Using the most difficult matrices kept the degree of difficulty of the task constant

frequent feedback enhanced the participants' learning about their ability and sought to reduce potential misestimations. Second, it allowed me to elicit the participants' beliefs about the number of matrices they thought they would be able to complete in the next round. These beliefs were elicited in each round after feedback was given. Third, it provided subjects with rest to minimize depletion.

As mentioned in the previous paragraph, subjects' beliefs about the number of matrices they expect to solve correctly in the next round were elicited. These beliefs correspond to the expression $\psi(e, \theta) = \mathbb{E}_1(\theta_i e)$ in the model. I did not provide monetary incentives in exchange of accurate beliefs. There are multiple reasons for this choice. First, Trautmann and van de Kuilen (2015) show that some desirable properties of incentive compatible methods can be achieved through non-incentivized belief elicitations. Second, I am interested in identifying belief differences due to the treatment assignment rather than in the accuracy of the elicited beliefs. Noise and belief inaccuracies stemming from participants' overconfidence or social desirability must be equally distributed across treatments due to randomization and thus potential belief differences must be due to effect of social status. This is indeed what the data analysis presented in Section 5 suggests. Third, to more closely relate to the survey evidence presented in Section 3, I asked participants to state a belief about their own performance on the task and, as in the survey, such belief does not have monetary consequences when (in)accurate.

Finally, for exactly half of the sessions, chosen at random, participants had the possibility of accessing a list showing how their current performance on the task ranked relative to that of the rest of participants in the same session. This feature was only available in Part 2 of the experiment. The ranking was determined by the participants' performance on Part 2 at the exact moment when it was accessed. To access it, participants had to click a button located at the right-bottom part of their screen. Looking at the ranking was costless and the program ensured that it could be quickly done. This attribute was included to investigate whether the treatment assignment affected the participants' decision to acquire additional information about their ability. If being exposed to more information about ability impedes the formation of motivated beliefs, then we should observe less access to this ranking among participants who benefit the most from those beliefs.

### 4.2. Experiment 2

The experiment was conducted at Tilburg University's CentERLAB. This experiment was conducted during the same weeks, but not in the same sessions, as Experiment 1. Participants were all students at the university and were recruited through an online system. The data consist of 8 sessions with a total of 138 subjects. On average a session lasted approximately

60 minutes. Between 11 and 23 participants took part in a session. The currency used in the experiment was euros. I used Z-Tree (Fischbacher, 2007) to implement and run the experiment. Participants earned on average 11.8 euros. The instructions of the experiment are presented in Appendix F.

The design of Experiment 2 had only one difference with respect to Experiment 1, namely that the assignment to the high status was determined by the participants' performance in Part 1. Specifically, participants with higher performance than at least half of the participants in the same session, were given the high status. The other half were assigned Low Status. In terms of the model, this treatment assignment implies $\epsilon = 1$.

There were several reasons behind implementing this experiment. First, it allows me to examine the robustness of potential findings in Experiment 1. If treatment effects are found in Experiment 1 and they are in the direction predicted by Proposition 1, then it must be that low ability participants in Experiment 2 exhibit low performance and have low beliefs as compared to high ability participants. That is because, according to the theory, low ability participants form motivated beliefs and attain high performance only if assigned to High Status. Second, these participants serve as an additional control group. Findings in Experiment 1 can be further corroborated by comparing the performance and beliefs levels of treated groups in Experiment 1 to participants with similar ability in Experiment 2.

### 4.3. Predictions

This subsection presents a set of predictions about the participants' performance, beliefs, and rank-access behavior in Experiment 1. They are based on the theoretical model of Section 2. The first two predictions are based on the equilibrium proved by Proposition 1 and regard the participants' performance and beliefs.

**Prediction 1.** *Performance is lowest for low ability subjects assigned to Low Status, higher for low ability subjects assigned to High Status, and is highest and comparable for high ability participants in both treatments.*

**Prediction 2.** *Beliefs are lowest for low ability subjects assigned to Low Status, higher for low ability subjects assigned to High Status, and are highest and comparable for high ability participants in both treatments.*

A validation of predictions 1 and 2 would contradict Proposition 2 and the result of the model's benchmark, as both results predict that the experimental treatments do not affect performance nor beliefs among low ability subjects. Also, these results would disregard an alternative prediction in which participants are completely naive and take status as informative signals about ability. That is mainly because in Predictions 1 and 2 high ability individuals do not behave in that way.

The remaining predictions regard the participants' access to the relative performance ranking. According to Corollary 1 and Corollary 2, if participants learn that the treatment assignment implied $\epsilon = 0$, the formation of motivated beliefs would be strongly constrained. Since the ranking provided accurate information about ability, it could uncover the relationship between status and ability and thus restrict the formation of motivated beliefs. As a consequence, participants who benefit the most from these beliefs, i.e. those with high status and low ability, should access this ranking less often.

**Prediction 3.** *Rank-access frequency is the lowest for low ability participants assigned to High Status.*

Another implication of Corollary 1 is that motivated beliefs, and the performance differences that they generate, can emerge despite participants having accessed the rank often enough to learn that $\epsilon = 0$. Moreover, Corollary 2 implies that these beliefs and performance differences are more likely to manifest the larger the magnitude of $\epsilon$ is, or, according to the alternative intuition given for that corollary, the larger participants believe the value of that parameter is. Since the ranking facilitates that participants learn that $\epsilon = 0$, less frequent access to the ranking facilitates subjects believing that $\epsilon$ is larger. Hence, the ease to form motivated beliefs when the rank is not accessed should be reflected on the magnitude of the treatment differences in beliefs and performance.

**Prediction 4.** *The performance and beliefs differences specified in Prediction 1 and Prediction 2 emerge regardless of whether the relative performance ranking was accessed or not. However, these differences become larger the less frequent the access to the rank is.*

The validity of these predictions will be evaluated in light of the experimental data analyzed in the next sections.

## 5. Experimental results

### 5.1. Performance

This section focuses on the data of Experiment 1. These data enable the evaluation of the effect of social status on performance, beliefs, and rank-access frequency through comparisons of these measures between participants with similar ability but assigned to different treatments. Where necessary, I also report the most relevant results of Experiment 2. Readers interested in the complete analysis of Experiment 2 are referred to Appendix E.

To closely link the results of Experiment 1 to the theoretical model, I classify participants into high ability, corresponding to $\theta_H$ in the model, and low ability, $\theta_L$. A participant is classified to have high ability if he correctly completes more matrices in Part 1 than at least half of the participants in the same session. Failing to be of high ability implied being classified as having low ability. Participants are not aware of this classification during the experiment. This classification entails considerable ability differences among these groups since high ability participants outperform low ability participants by 2.21 standard deviations in Part 1, a significant difference (Hedge's g, $p < 0.001$).

A successful treatment randomization must guarantee that participants' ability is comparable across treatments. I find no difference in average performance in Part 1 between low ability participants assigned to the different treatments ($U = 0.028, p = 0.977$). However, high ability participants assigned to the low status exhibit higher performance in Part 1 than high ability participants assigned the high status ($U = 1.992, p = 0.041$). This difference is taken into account in what remains of the analysis, but posses no threat to the validity of the main results of the paper.

We are now in a position to analyze the participants' performance in Part 2. Table 3 presents the descriptive statistics of participants' performance by treatment and ability. The main finding is that low ability participants exhibit higher performance when assigned the high status ($U = 2.020, p = 0.029$).[18] The effect size of this difference is 0.53 standard deviations (Hedge's g, $p = 0.01$, with 1000 bootstrap replications).[19] The average performance level achieved by low ability participants assigned to High Status is comparable to that achieved by high ability participants assigned to the same treatment ($U = 0.658, p = 0.510$). Additionally, low ability participants assigned to the low status treatment are outperformed by high ability participants assigned the same treatment ($U = 2.378, p < 0.01$). Finally, I find no empirical evidence of a difference in performance between high ability participants assigned to the different treatments ($U = 0.913, p = 0.36$).[20]

These differences in performance constitute the first result of the paper and validate Prediction 1.

[18]Unless otherwise specified, I use the Wilcoxon Mann-Whitney test for pairwise comparisons and report the standardized U statistic and the respective p-value.

[19]The statistical power of this test is $1 - \beta = 0.73$ at the 5 % significance level.

[20]Since high ability participants exhibit performance differences across the treatments in Part 1, the latter finding could be interpreted as high ability participants displaying higher performance when assigned to High Status. The analysis using regressions presented in Appendix D.2 corroborate the result that high ability participants assigned to different treatments exhibit similar average performance. This result also emerges when average ability on the task in Part 1 and other control variables are included in the regressions.

Table 3: Descriptive statistics of performance in Part 2 of Experiment 1

| Type/Treatment | High Status | Low Status | Total |
|---|---|---|---|
| High Ability | 22.285 | 24.658 | 23.695 |
| | (7.215) | (8.676) | (8.144) |
| Low Ability | 24.771 | 19.621 | 22.437 |
| | (11.476) | (6.630) | (9.863) |
| Total | 23.68 | 22.65 | 23.139 |
| | (9.744) | (8.224) | (8.952) |

Note: This table presents the averages and standard deviations of the performance in the second part of Experiment 1 by experimental treatment and subject ability. Standard deviations are presented in parentheses.

**Result 1.** *Average performance is lowest for Low ability participants assigned Low Status and highest and comparable for high ability participants and Low ability participants with High Status.*

This result is in line previous experimental research investigating the effect of social status on economic behavior (Ball and Eckel, 1998, Ball et al., 2001, Eckel and Wilson, 2007, Kumru and Vesterlund, 2010, Eckel et al., 2010). When assigned the high status, participants with low ability on the task derived higher average earnings as a result of exerting higher effort and attaining higher performance on the task.

### 5.2. Beliefs

Next, I examine the influence of the treatments on beliefs. The analysis focuses on the *sum* of the participants' beliefs over all rounds in Part 2. Table 4 presents the descriptive statistics of this measure of beliefs by treatment and by ability. While high ability participants exhibit similar average beliefs across treatments ($U = 0.177, p = 0.859$), low ability participants exhibit higher average beliefs when assigned to High Status ($U = 2.071, p = 0.038$). In fact, the average belief level exhibited by low ability participants with the high status is comparable to that of high ability participants assigned the same treatment ($U = 0.672, p = 0.501$). Hence, the assignment to High Status induced higher beliefs among low ability participants.

These differences in beliefs constitute the second result of the paper and confirm Prediction 2.

**Result 2.** *Average beliefs are lowest for Low ability participants assigned Low Status and highest comparable for high ability participants and Low ability participants with High Status.*

This result corroborates a conjecture posed by Ball and Eckel (1998) (pp. 506 and pp. 510)

Table 4: Performance beliefs in Part 2 of Experiment 1

| Ability / Treatment | High Status | Low Status | Total |
|---|---|---|---|
| High Ability | 31.285 | 31.804 | 31.594 |
| | (8.944) | (8.721) | (8.793) |
| Low Ability | 29.8 | 27.620 | 28.81 |
| | (8.442) | (9.484) | (10.458) |
| Total | 32.161 | 29.559 | 31.691 |
| | (9.240) | (9.113) | (9.249) |

Note: This table presents the averages and standard deviations of aggregated beliefs in the second part of Experiment 1 by experimental treatment, and subject ability. Standard deviations are presented in parentheses.

and Ball et al. (2001)(pp.169): expectations about reasonable economic outcomes are changed by the social status assignment. In the context of this experiment, status changed the low ability individuals' expectation about their own abilities and thus changed the performance levels that they considered to be reasonable.

A set of robustness checks further corroborate Result 1 and Result 2. I perform regressions to evaluate the significance of these treatment effects while controlling for a set of relevant variables. The estimates presented in Table 12 of Appendix D.2 confirm all aforementioned findings. Appendix D.3 shows that comparable results are found when the binary classification of social status is replaced by a continuous measure of ability, i.e. performance in Part 1 of the experiment. Hence, these results are not artifacts of using a coarse classification of ability. Appendix E show that low ability participants exhibit lower average performance and beliefs as compared to high ability participants in Experiment 2. This result holds when the binary measure of ability is replaced by a continuous measure. Therefore, the treatment differences in Experiment 1 are due to the treatment assignment. Additionally, the aforementioned findings are robust to treatment and control comparisons across the two experiments.[21] Appendix D.4 presents an alternative analysis of the data that focuses on the participants' beliefs *in a given round*. I find that the treatment differences in beliefs appear as of the second round. Hence, Result 2 is not due to misinformation or incorrect priors that participants might hold about their ability. Instead, it materializes after participants were exposed to feedback and after they were acquainted with the task.

An alternative explanation to Result 1 and Result 2 is that they are due to experimenter

[21] Specifically, low ability subjects assigned to High Status in Experiment 1 exhibit higher average performance than low ability subjects in Low Status in Experiment 1 and Experiment 2 ($U = 1.894, p = 0.028$). Also, I find no significant differences in performance between high ability subjects assigned to low status in Experiment 1 and high ability subjects in High Status in both experiments ($U = 0.466p = 0.6411$).

demand. That is, these treatment differences are due to participants inferring the objective of the experiment and reacting to the status assignment accordingly. In a recent paper, De Quidt et al. (2018) show that these effects are modest. In their closest task, an incentivized real-effort task, they found that the experimenter demand effect has a magnitude of 0.07 standard deviations. Hence, *if* the treatment protocol triggered an experimenter demand effect, the effect size of the treatment effects would remain significant after that magnitude is subtracted.[22] Furthermore, it is important to emphasize that the fact that the status manipulation lead participants to believe that one group is ranked higher than another is not an experimental demand effect in the usual sense—an accidental effect of what the experimenter wants the result to be—but a deliberate manipulation belonging to a treatment variable (See the discussion in Ball et al. (2001) footnote 8).

Finally, I study how beliefs relate to performance by analyzing the empirical properties of the difference between these two variables. I find a generalized tendency of participants to exhibit overconfident beliefs: participants stated average beliefs that were 7.09 matrices higher than their average performance level. This imprecision is similar across treatments and across ability, indicating that the randomization guaranteed similar degrees of overconfidence across these groups. For instance, low ability participants, who exhibited belief differences across treatments, exhibit an average gap between performance and beliefs of $-8$ when assigned Low Status and an average gap of $-5.02$ when assigned High Status, differences that are statistically indistinguishable ($U = -1.109, p = 0.271$).[23] Therefore, Result 2 is not an artifact of an unbalanced differences in belief inaccuracy among groups, but is treatment difference resulting from comparing groups with similar degrees of overconfidence.

### 5.3. Access to relative performance feedback

I use the data on the participants' rank access behavior to investigate Prediction 3 and Prediction 4. These predictions go to the heart of the proposed mechanism. If subjects strategically misinterpret social status to generate motivated beliefs, then we should observe differences in rank-access frequency across the treatments. That is because low ability subjects are sophisticated and expect to rank low and this information hampers the formation of motivated beliefs.

I first focus on the frequency at which participants accessed the rank and whether being

[22] The average performance between the treatments accounting for potential experimenter demand effects is of 0.46 standard deviations with a bootstrapped confidence interval of $[0.88, 0.06]$

[23] Among high ability participants the average gap is $-9$ for subjects in High Status and $-7.146$ for subjects in Low Status, a non-significant difference ($U = 0.514, p = 0.607$). Among subjects assigned the high status treatment, I find no significant difference in this gap between high and low ability subjects ($U = 1.066, p = 0.286$). The same conclusion is reached when the performance-beliefs gap of participants with different ability but assigned to the low status treatment is compared ($U = 0.036, p = 0.9714$).

assigned to High Status influenced this decision. Table 5 presents descriptive statistics of rank-access frequency measured at different points in time during Part 2. Panel A, Panel B, and Panel C present the descriptive statistics of rank-access frequency after the first round was over, after the first three rounds were over, and at the end of the experiment, respectively. The reason for using these different measurements is the great heterogeneity in rank access over rounds, which, as it will become evident, affects the results.

Table 5 shows that participants who had the chance to access the ranking did so 2.66 times on average. A comparison across panels evidences that the frequency at which the rank was accessed considerably increased over rounds. Specifically, participants accessed the rank 0.36 times on average after the first round, and this number surged to 1.055 after the third round. Such increase in rank-access frequency over rounds is explained by a higher proportion of participants accessing the rank at least once. After round 1 approximately 18% of participants who had access to the rank accessed it at least once. This proportion increased to 38% at the end of round 3, and further increased to 63% at the end of the experiment.

I do not find differences in rank-access frequency between participants assigned to the different treatments ($U = 0.758, p = 0.448$), or between participants with different ability ($U = 0.774, p = 0.439$). These conclusions are robust to using the different rank-access measures. To evaluate whether participants with similar ability but who were assigned to different treatments exhibit differences in rank-access frequency, I perform regressions. The main rationale for using regressions rather than pairwise tests is that the latter tests can be under-powered due to low number of participants belonging to a treatment, having some ability in the task, *and* that had access to the ranking.

Table 6 presents the estimates of regressions of rank-access frequency on a treatment dummy, an ability dummy, interactions between these two variables, and, for some specifications, control variables. Low ability participants in High Status display lower access-frequency than participants with similar ability in Low Status. This result holds for the rank-access measures taken after the first round and after the first three rounds in Part 2.[24] The rank-access frequency of low ability participants in the high status is also lower than that displayed by high ability participants ($\chi^2(1) = 4.95, p = 0.026$). This result confirms Prediction 3 and demonstrates that subjects benefiting the most from believing $\epsilon > 0$ were less likely to acquire information that could contradict these beliefs. Appendix E further corroborates this result by showing that low ability and high ability participants in Experiment 2 exhibit similar rank-access frequency.

This result is further validated by the participants' answers to the post-experimental

[24]The joint tests supporting these conclusions yield ($\chi^2(1) = 5.61, p = 0.017$) for the estimates presented in column 2 of Table 6 and ($\chi^2(1) = 4.67, p = 0.030$) for the estimates column 4 of Table 6.

Table 5: Rank-access frequency in Experiment 1 by treatment and by ability

| Ability/Treatment | High Status | Low Status | Total |
|---|---|---|---|
| Panel A: First round | | | |
| High Ability | 0.727 | 0.181 | 0.363 |
| | (7.215) | (1.272) | (1.025) |
| Low Ability | 0.2 | 0.312 | 0.25 |
| | (0.523) | (0.602) | (0.554) |
| Total | 0.387 | 0.236 | .304 |
| | (0.882) | (0.751) | (0.809) |
| Panel B: First three rounds | | | |
| High Ability | 1.727 | 0.636 | 1 |
| | (3.289) | (1.705) | (2.358) |
| Low Ability | 0.95 | 1.187 | 1.055 |
| | (1.959) | (1.558) | (1.771) |
| Total | 1.225 | 0.868 | 1.028 |
| | (2.486) | (1.646) | (2.057) |
| Panel C: All rounds | | | |
| High Ability | 3.363 | 2.227 | 2.606 |
| | (4.884) | (3.624) | (4.046) |
| Low Ability | 3 | 2.375 | 2.722 |
| | (3.128) | (2.446) | (2.824) |
| Total | 3.129 | 2.289 | 2.66 |
| | (3.766) | (9.113) | (3.437) |

Note: This table presents the averages and standard deviations of rank-access frequency in the second part of Experiment 1 by experimental treatment, and subject ability. Standard deviations are presented in parentheses.

survey questions. In that survey subjects where asked if the assignment to the high status was deserved, due to their performance, and if it was fair. The questions can be found in Appendix F. Table 16 in Appendix D.5 shows that low ability participants assigned to the high status are more likely to state that the medal is deserved, fair, and due to their performance as compared to subjects with similar ability but assigned to the low status. Hence, when advantageous, subjects sought to maintain the belief that $\epsilon > 0$.

Table 6: Determinants of rank-access in Experiment 1

| | (1) Times Rank Round 1 | (2) Times Rank Round 1 | (3) Times Rank Round 1-3 | (4) Times Rank Round 1-3 | (5) Times Rank All rounds | (6) Times Rank All rounds |
|---|---|---|---|---|---|---|
| Low Ability | 0.506 | 0.796* | 0.746 | 1.226* | 0.351 | 1.348 |
| | (0.322) | (0.382) | (0.478) | (0.655) | (1.008) | (1.204) |
| High Status | 0.676* | 0.647** | 0.651 | 0.601 | 0.278 | 0.168 |
| | (0.364) | (0.312) | (0.636) | (0.690) | (1.397) | (1.784) |
| Low Ability* High Status | -0.834* | -0.945*** | -1.452* | -1.551* | 0.095 | -0.057 |
| | (0.434) | (0.346) | (0.890) | (0.900) | (1.737) | (1.966) |
| Group Size | | 0.002 | | -0.035 | | -0.127 |
| | | (0.037) | | (0.096) | | (0.210) |
| Fair Medal | | 0.220 | | 0.120 | | -0.328 |
| | | (0.254) | | (0.485) | | (0.979) |
| Locus of Control | | -0.166*** | | -0.360* | | -0.122 |
| | | (0.059) | | (0.189) | | (0.351) |
| Female | | -0.425** | | -0.879* | | -2.149** |
| | | (0.205) | | (0.533) | | (1.016) |
| Belief practice round | | -0.036* | | -0.042 | | -0.065 |
| | | (0.021) | | (0.115) | | (0.237) |
| Task performed before | | 0.087 | | -0.405 | | -0.461 |
| | | (0.225) | | (0.723) | | (1.797) |
| $ln(\delta)$ | 0.221 | -0.104 | 1.116*** | 0.995*** | 1.405*** | 1.246*** |
| | (0.564) | (0.730) | (0.375) | (1.140) | (0.286) | (0.344) |
| Observations | 69 | 69 | 69 | 69 | 69 | 69 |
| Log Likelihood | -43.297 | -39.236 | -89.620 | -86.691 | -146.453 | -142.691 |

Note: This table presents marginal effects of negative binomial regressions of the model $TimesRank_i = \beta_0 + \beta_1 LowAbility + \beta_2 HighStatus + \beta_3 LowAbility * HighStatus + Controls'\Gamma + \epsilon_i$, with $\epsilon \sim poisson(\lambda)$. "Times Rank" is the frequency of rank-access by a subject in the second part of Experiment 1,"High Status" is a dummy variable that captures whether the subject was assigned to the high status treatment "Low Ability" is a dummy variable that captures whether the subject was classified as having low ability on the task. $ln(\delta)$ is the estimated dispersion from the mean. Standard errors presented in parentheses. *** denotes significance at the 0.01 level, ** denotes significance at the 0.05 level, * denotes significance at the 0.1 level.

Table 6 also shows that when the rank-access frequency is examined after all rounds elapsed, these treatment differences dissipate.[25] This finding, along with the fact that there is a significantly higher proportion of participants accessing the rank in the last rounds of the experiment, suggest an end-of-the-game effect. That is, in the last rounds when there is little time left and small margin to increase earnings, the majority of participants access the ranking to check how well they did in the experiment.

The emergence of such end-of-the-game effect confines the empirical validity of Prediction 3 to all but the last rounds of Part 2.

**Result 3.** *Low ability participants display lower access to the rank when assigned to High status. This result dissipates in the last two rounds of Part 2.*

Finally, I investigate if Prediction 4 holds. Table 7 presents the estimates of statistical models similar to those presented in Table 12 with the difference that "Times Rank", the variable that captures the frequency of rank-access, is interacted with the treatment dummy and the ability dummy. I find that treatment effects are larger among participants with low ability who did not access the rank ($\chi^2(1) = 5.61, p = 0.017$). A result that is robust to the used measure of rank-access frequency. This finding justifies the differences in rank access documented in Result 3. Furthermore, the estimates in columns (3)-(6) of Table 7 show that low ability participants, regardless of whether they access or not the rank, exhibit treatment effects. Note that the significance of this finding is sensitive to the choice of rank-access frequency measures. In particular, treatment effects are found when rank-access frequency is measured after the third round and after all rounds have elapsed. I conjecture that this is due to the low number of participants who accessed the rank in the first round (9% of participants), which impedes the precise estimation of treatment effect estimates.

These results validate Prediction 4 for all but the first two rounds of the experiment, when treatment effects can be precisely estimated.

**Result 4.** *Low ability participants display stronger treatment effects the less they access the rank. Treatment effects are found regardless of whether these participants access the rank, this latter result emerges after the first three rounds of Part 2.*

[25] Low ability participants assigned high status exhibit similar average rank-access as compared to low ability subjects assigned the low status ($\chi^2(1) = 0.05, p = 0.8171$). Also, low ability subjects with low status exhibit similar average rank-access frequency as compared to high ability subjects assigned to the high status ($\chi^2(1) = 0.42, p = 0.516$).

Table 7: Treatment effects and rank-access in Experiment 1

| | (1) Performance | (2) Beliefs | (3) Performance | (4) Beliefs | (5) Performance | (6) Beliefs |
|---|---|---|---|---|---|---|
| Low Ability | -5.873*** | -4.820* | -4.787** | -3.701** | -3.899* | -2.038 |
| | (2.104) | (2.484) | (2.160) | (1.568) | (2.159) | (2.494) |
| High Status | -1.770 | 0.459 | -1.682 | -0.594 | -1.138 | 0.765 |
| | (1.820) | (1.935) | (1.886) | (1.412) | (1.967) | (2.013) |
| Low Ability * High Status | 7.132** | 1.674 | 6.139** | 3.388 | 4.111 | -2.268 |
| | (2.926) | (3.187) | (2.938) | (2.096) | (2.864) | (3.279) |
| Low Ability * Times Rank | -0.474 | -2.076 | -2.082** | -2.097** | -1.581*** | -2.609** |
| | (2.207) | (4.107) | (1.059) | (0.897) | (0.589) | (1.124) |
| High Status* Times Rank | -2.754 | -0.884 | -1.124 | -0.689 | -0.799 | -0.445 |
| | (2.342) | (1.326) | (0.750) | (0.440) | (0.499) | (0.506) |
| High Status* Low Ability Times Rank | -1.517 | 6.185 | 1.607 | 2.635*** | 2.206** | 3.819*** |
| | (3.359) | (4.312) | (1.059) | (0.897) | (0.882) | (1.212) |
| Times Rank | 2.106*** | -1.893** | 1.054** | 0.073 | 0.628*** | -0.004 |
| | (0.789) | (0.953) | (0.419) | (0.394) | (0.182) | (0.446) |
| Group Size | -0.569 | -0.023 | -0.601* | 0.098 | -0.605* | 0.049 |
| | (0.359) | (0.432) | (0.354) | (0.309) | (0.347) | (0.416) |
| Fair Medal | -1.988 | -2.205 | -2.075 | -0.668 | -2.164 | -2.266 |
| | (1.801) | (1.856) | (1.854) | (1.205) | (1.878) | (1.717) |
| Locus of Control | 0.650 | 0.252 | 0.742* | 0.351 | 0.659 | 0.497 |
| | (0.439) | (0.547) | (0.434) | (0.369) | (0.423) | (0.494) |
| Female | 3.031* | -0.340 | 3.091* | -0.694 | 3.323* | -0.883 |
| | (1.662) | (2.075) | (1.720) | (1.348) | (1.901) | (1.876) |
| Belief practice round | -0.049 | 0.358 | -0.018 | 1.442** | -0.020 | 0.459* |
| | (0.197) | (0.273) | (0.197) | (0.139) | (0.199) | (0.260) |
| Task Performed before | 5.914*** | 7.574** | 5.718*** | 5.556*** | 5.637*** | 6.238** |
| | (2.132) | (3.024) | (2.108) | (1.715) | (1.941) | (2.934) |
| $ln(\delta)$ | 0.387 | 0.368 | 0.387 | 0.264 | 0.295 | 0.264 |
| | (0.308) | (0.249) | (0.314) | (0.249) | (0.287) | (0.227) |
| Times Rank in | Round 1 | Round 1 | Round 1-3 | Round 1-3 | All rounds | All rounds |
| Observations | 133 | 133 | 133 | 133 | 133 | 133 |
| Log Likelihood | -452.900 | -473.220 | -452.872 | -470.468 | -451.194 | -468.945 |

Note: This table presents marginal effects of negative binomial regressions of the model $Performance_i = \beta_0 + \beta_1 LowAbility + \beta_2 HighStatus + \beta_3 LowAbility * HighStatus + \beta_4 LowAbility * TimesRank + \beta_5 HighStatus * TimesRank + \beta_6 LowAbility * HighStatus * TimesRank + Controls'\Gamma + \epsilon_i$, with $\epsilon \sim poisson(\lambda)$. "Performance" is the number of correctly solved matrices in the second stage of Experiment 1,"Times Rank" is the occurrence of rank-access by a subject in the second part of Experiment 1,"High Status" is a dummy variable that captures whether the subject was assigned to the high status treatment "Low Ability" is a dummy variable that captures whether the subject was classified as having low ability on the task. $ln(\delta)$ is the estimated dispersion from the mean. Standard errors presented in parentheses. *** denotes significance at the 0.01 level, ** denotes significance at the 0.05 level, * denotes significance at the 0.1 level.

## 6. Conclusion and Discussion

This paper demonstrated that internal constraints generated by social status influence the achievement of economic outcomes. As a consequence, social standings affect economic success not only through the pecuniary advantages that they encompass but also through the beliefs they trigger in individuals. The main economic implication of this paper is that societies with institutions that instigate acute social differences are stimulating economic disparities through unexpected channels. For instance, members of historically discriminated groups are not having the opportunity to form beneficial and encouraging beliefs, making their disadvantage self-fulfilling.

A theoretical framework formalized the proposed mechanism and provided a set of testable predictions. According to the model, individuals incorrectly but deliberately use a high social status assignment to boost their beliefs, which in turn allows them to achieve outcomes that otherwise would have been too costly. This generates a disadvantage for low status individuals who cannot access these self-serving beliefs and, as a result, achieve lower economic outcomes. In this way, a seemingly inoffensive status disparity becomes harmful and self-sustaining.

Two analyses empirically validated the theoretical predictions and complemented each other. The survey data analysis demonstrated that the proposed mechanism materializes outside the laboratory and quantified the extent to which it affected the lives of individuals. I find that individuals born in a higher status household were more likely to exhibit higher educational aspirations and, consequently, were more likely to achieve a higher status during their adulthood. The laboratory data analysis provided clean causal evidence of the mechanism that did not require the strong econometric assumptions imposed in the survey data analysis. Participants with similar ability on a task exhibited significantly higher performance and beliefs when assigned to the high status.

To address the disparities proved in this paper, it would be unrealistic to formulate policies aimed at raising everyone's perceived social position since social status is a zero-sum concept. However, macroeconomic policies that seek to reduce wealth inequalities can be useful in relaxing the internal constraints generated by acute social status differences. Furthermore, policies designed to provide disadvantaged individuals with the opportunity to develop relevant abilities can help them attain high economic outcomes. That is because model and the experiment show that high-ability individuals do not experience the adverse psychological effects from belonging to a low social status. Moreover, policies with the goal of shaping the way in which individuals form beliefs could help them attain high outcomes. For instance, programs that stimulate ambitious aspirations by presenting the stories of successful role models that share similar backgrounds, or programs that facilitate the access to therapy aimed at improving self-confidence and self-esteem.

## References


Auriol, E. and Renault, R. (2008). Status and Incentives. *The RAND Journal of Economics*, 39(1):305–326.

Ball, S. and Eckel, C. C. (1998). The economic value of status. *The Journal of Socio-Economics*, 27(4):495–514.

Ball, S., Eckel, C. C., Grossman, P. J., and Zame, W. (2001). Status in markets. *Quarterly Journal of Economics,*, 116(1):161–188.

Benabou, R. (2015). The Economics of Motivated Beliefs. *Revue d'Economie Politique*, 125(5):665–685.

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

Bernard, T., Dercon, S., Orkin, K., and Taffesse, A. (2014). *The future in mind: Aspirations and forward-looking behaviour in rural Ethiopia*. Centre for Economic Policy Research London.

Besley, T. (2016). Aspirations and the political economy of inequality. *Oxford Economic Papers*, 69:1–35.

Bharadwaj, P., Løken, K. V., and Neilson, C. (2013). Early life health interventions and academic achievement. *American Economic Review*, 103(5):1862–1891.

Blanden, J., Gregg, P., and Macmillan, L. (2007). Accounting for Intergenerational Income Persistence: Non-Cognitive Skills, Ability and Education. *Economic Journal*, 117:43–60.

Bogliacino, F. and Ortoleva, P. (2015). The Behavior of Others as a Reference Point. *Documentos FCE-CID Escuela de Economía*, 67:491, 502.

Borjas, G. J. (1995). The Economic Benefits from Immigration. *Journal of Economic Perspectives*, 9(2):3–22.

Bors, D. A. and Stokes, T. L. (1998). Raven's advanced progressive matrices: Norms for first-year university students and the development of a short form. *Educational and Psychological Measurement*, 58:382–398.

Bowles, S., Gintis, H., and Osborne, M. (2001). The Determinants of Earnings: A Behavioral Approach. *Journal of Economic Literature*, 39(4):1137–1176.

Charness, G., Masclet, D., and Villeval, M. C. (2014). The Dark Side of Competition for Status. *Management Science*, 60(June 2015):38–55.

Chetty, R. and Hendren, N. (2018). The impacts of neighborhoods on intergenrational mobility I : childhood exposure effects. *The Quarterly Journal of Economics*, 133(3):1107–1162.

Chetty, R., Hendren, N., and Katz, L. F. (2015). The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment. *American Economic*

*Review*, 106(4):855–902.

Chiapa, C., Garrido, J. L., and Prina, S. (2012). The effect of social programs and exposure to professionals on the educational aspirations of the poor. *Economics of Education Review*, 31(5):778–798.

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

Cunha, F. C., Heckman, J. J., and Schennach, S. M. (2010). Estimating the Technology of Cognitive and Noncognitive Skill Formation. *Econometrica*, 78(3):883–931.

Dalton, P. S., Ghosal, S., and Mani, A. (2016). Poverty and Aspirations Failure. *Economic Journal*, 126:165–188.

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

Eckel, C. C., Fatas, E., and Wilson, R. K. (2010). Cooperation and Status in Organizations. *Journal of Public Economic Theory*, 12(4):737–762.

Eckel, C. C. and Wilson, R. K. (2007). Social learning in coordination games: does status matter? *Experimental Economics*, 10(3):317–329.

Eil, D. and Rao, J. (2011). The Good News-Bad News Effect: Asymmetric Processing of Objective Information about Yourself. *American Economic Journal: Microeconomics*, 03(02):114–138.

Field, E., Robles, O., and Torero, M. (2009). Iodine deficiency and schooling attainment in Tanzania. *American Economic Journal: Applied Economics*, 1(4):140–69.

Fischbacher, U. (2007). Z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10:171–178.

Genicot, G. and Ray, D. (2017). Aspirations and Inequality. *Econometrica*, 85(2):489–519.

Ghatak, M. (2015). Theories of poverty traps and anti-poverty policies. *World Bank Economic Review*, 29:S77–S105.

Heckman, J. J. (1981). Heterogeneity and state dependence. In Rosen, S., editor, *Studies in Labor Markets*, chapter 3, pages 91–139. University of Chicago Press.

Heckman, J. J., Stixrud, J., and Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics*, 24(3):411–482.

Hopkins, E. and Kornienko, T. (2006). Inequality and growth in the presence of competition for status. *Economics Letters*, 93(2):291–296.

Janzen, S. A., Magnan, N., Sharma, S., and Thompson, W. M. (2017). Aspirations failure and formation in rural nepal. *Journal of Economic Behavior & Organization*, 139:1–25.

Jensen, R. (2010). The (perceived) returns to education and the demand for schooling. *Quarterly Journal of Economics*, 125(2):515–548.

Kamenica, E. and Gentzkow, M. (2011). Bayesian persuasion. *American Economic Review*, 101(6):2590–2615.

Koszegi, B. (2006). Ego Utility, Overconfidence, and Task Choice. *Journal of the European Economic*

*Association*, 4(June):673–707.

Kuhnen, C. M. and Tymula, A. (2012). Feedback, Self-Esteem, and Performance in Organizations. *Management Science*, 58(June 2015):94–113.

Kumru, C. S. and Vesterlund, L. (2010). The Effect of Status on Charitable Giving. *Journal of Public Economic Theory*, 12(4):709–735.

Laajaj, R. (2017). Endogenous time horizon and behavioral poverty trap: Theory and evidence from mozambique. *Journal of Development Economics*, 127:187–208.

Macours, K. and Vakis, R. (2014). Changing households' investment behaviour through social interactions with local leaders: Evidence from a randomised transfer programme. *The Economic Journal*, 124(576):607–633.

Mobius, M., Niederle, M., Paul, N., and Rosenblat, T. (2014). Managing Self-Confidence. *mimeo*.

Moldovanu, B., Sela, A., and Shi, X. (2007). Contests for Status. *Journal of Political Economy*, 115(2):338–363.

Plug, E. and Vijverberg, W. (2003). Schooling , Family Background , and Adoption : Is It Nature or Is It Nurture? *Journal of Political Economy*, 111(3):611–641.

Plug, E. and Vijverberg, W. (2005). Does Family Income Matter for Schooling Outcomes ? Using Adoptees as a Natural Experiment. *Economic Journal*, 506:879–906.

Raven, J. (1989). The Raven Progressive Matrices : A Review of National Norming Studies and Ethnic and Socioeconomic Variation within the United States. *Journal of Educational Measurement*, 26(1):1–16.

Ridgeway, C. L. and Walker, H. A. (1995). Status structures. *Sociological perspectives on social psycholoy*, pages 281–310.

Ross, P. H. (2019). Occupation aspirations, education investment, and cognitive outcomes: Evidence from indian adolescents. *World Development*, 123:104613.

Ross, P. H., Glewwe, P., Prudencio, D., and Wydick, B. (2021). Developing educational and vocational aspirations through international child sponsorship: Evidence from kenya, indonesia, and mexico. *World Development*, 140:105336.

Sacerdote, B. (2002). The nature and nurture of economic outcomes. *American Economic Review*, 92(2):344–348.

Trautmann, S. T. and van de Kuilen, G. (2015). Belief Elicitation: A Horse Race among Truth Serums. *Economic Journal*, 125(589):2116–2135.

# Appendix A: Proofs

## Preliminaries

Before providing the proofs of Propositions 1 and 2, I characterize the reaction function of Self 1 in Lemma 1. The merit of that Lemma is that it proves the existence of a threshold probability $\hat{\lambda}_{-k}$ that makes Self 1 indifferent between choosing $e_H$ or $e_L$ upon receiving a signal $\tilde{\theta}_k$, where $k = i$ or $k = j$ and $i \neq j$.

**Lemma 1.** *Under Assumptions 1, 2, and 3, the optimal effort chosen by Self 1 after receiving a signal $\tilde{\theta}_k$ is given by the function:*

$$e(\tilde{\theta}_k) = \begin{cases} e_H & \textit{if } \tilde{\theta}_H \textit{ and } \lambda_L \in [\hat{\lambda}_L, 1], \\ e_H & \textit{if } \tilde{\theta}_L \textit{ and } \lambda_H \in [0, \hat{\lambda}_H]. \end{cases}$$

*Where $\hat{\lambda}_L$ and $\hat{\lambda}_H$ are threshold probabilities satisfying:*

$$\hat{\lambda}_L = 1 - \frac{(\theta_H(a+s) - c)\Big(qr + \lambda_H q(1-r) + \epsilon(1-\lambda_H)\Big)}{\Big(c - \theta_L(a+s)\Big)\chi\left((1-q)r - \epsilon\right)},$$

*and*

$$\hat{\lambda}_H = 1 - \frac{(\theta_H(a+s) - c)\Big((1-q)(1-r) + \lambda_L(1-q)r + \epsilon(1-\lambda_L)\Big)}{\Big(c - \theta_L(a+s)\Big)\chi(q(1-r) - \epsilon)}.$$

*Proof.* Suppose that Self 1 receives $\tilde{\theta}_H$ from Self 0. Upon receiving this signal, he chooses $e_H$ as long as:

$$\mathbb{E}_1\Big(U(e_H, \theta_i)|\tilde{\theta}_H\Big) \geq \mathbb{E}_1\Big(U(e_L, \theta_i)|\tilde{\theta}_H\Big) \Leftrightarrow p_H(\lambda_H, \lambda_L)\theta_H(a+s) + \Big(1 - p_H(\lambda_H, \lambda_L)\Big)\theta_L(a+s) \geq c, \tag{9}$$

where $p_H(\lambda_H, \lambda_L)$ is defined in equation (4). Algebraic manipulations yield the following inequality:

$$\lambda_L \geq 1 - \frac{(\theta_H(a+s) - c)\Big(qr + \lambda_H q(1-r) + \epsilon(1-\lambda_H)\Big)}{\Big(c - \theta_L(a+s)\Big)\chi\left((1-q)r - \epsilon\right)}. \tag{10}$$

Thus, Self 1 chooses $e_H$ after receiving $\tilde{\theta}_H$ if $\lambda_L \geq \hat{\lambda}_L$ where

$$\hat{\lambda}_L := 1 - \frac{(\theta_H(a+s) - c)\left(qr + \lambda_H q(1-r) + \epsilon(1-\lambda_H)\right)}{\left(c - \theta_L(a+s)\right)\chi\left((1-q)r - \epsilon\right)}. \tag{11}$$

Next, suppose that Self 1 receives $\tilde{\theta}_L$. Choosing $e_H$ is profitable as long as :

$$\mathbb{E}_1\Big(U(e_H,\theta_i)|\tilde{\theta}_L\Big) \geq \mathbb{E}_1\Big(U(e_L,\theta_i)|\tilde{\theta}_L\Big) \Leftrightarrow p_L(\lambda_H,\lambda_L)\theta_H(a+s)+\Big(1-p_L(\lambda_H,\lambda_L)\Big)\theta_L(a+s) \geq c, \tag{12}$$

where $p_L(\lambda_H,\lambda_L)$ as in equation (5). Algebraic manipulations yield the following condition:

$$\lambda_H \leq 1 - \frac{(\theta_H(a+s)-c)\Big((1-q)(1-r)+\epsilon+\lambda_L\left((1-q)r-\epsilon\right)\Big)}{\Big(c-\theta_L(a+s)\Big)\chi(q(1-r)-\epsilon)} \tag{13}$$

Hence, $e_H$ is chosen by Self 1 after receiving $\tilde{\theta}_L$ if $\lambda_H \leq \hat{\lambda}_H$ where

$$\hat{\lambda}_H := 1 - \frac{(\theta_H(a+s)-c)\left((1-q)(1-r)+\epsilon+\lambda_L\left((1-q)r-\epsilon\right)\right)}{\left(c-\theta_L(a+s)\right)\chi(q(1-r)-\epsilon)}. \tag{14}$$

■

## Proofs of the main results

## Proposition 1

*Proof.* Suppose that Self 0 with $\theta_H$ chooses $\lambda_H = 1$. According to Lemma 1, Self 1 chooses $e_L$ after receiving $\tilde{\theta}_L$. Instead, after receiving $\tilde{\theta}_H$ he chooses either $e_H$ or $e_L$ depending on the value of $\lambda_L$. Suppose that $\lambda_L = 0$ induces $e_L$. For Self 0, this strategy is profitable if

$$\begin{aligned} \mathbb{E}_0\Big(U(\lambda_L=1,e_L)\Big) \leq& \mathbb{E}_0\Big(U(\lambda_L=0,e_L)\Big) \Leftrightarrow \\ \theta_L(a+s)e_L \leq& \theta_L(a+s)e_L + se_L(\theta_H-\theta_L)p_H(\lambda_H=1,\lambda_L=0) - m, \end{aligned} \tag{15}$$

a condition that can be rewritten as an upper bound of $m$:

$$m \leq \frac{se_L q(\theta_H-\theta_L)}{q+\chi\left(r(1-q)-\epsilon\right)}. \tag{16}$$

Suppose that $\lambda_L = 0$ induces $e_H$. This strategy is profitable for Self 0 if

$$\begin{aligned} \mathbb{E}_0\Big(U(\lambda_L=1,e_L)) \leq& \mathbb{E}_0(U(\lambda_L=0,e_H)\Big) \Leftrightarrow \\ \theta_L(a+s)e_L \leq& \theta_L(a+s)e_H + se_H(\theta_H-\theta_L)p_H(\lambda_H=1,\lambda_L=0) - m - c, \end{aligned} \tag{17}$$

which holds if $c \geq (a+s)\theta_L$, a condition that always holds, and if $m$ attains the upper bound:

$$m \leq \frac{se_H q(\theta_H-\theta_L)}{q+\chi\left(r(1-q)-\epsilon\right)}. \tag{18}$$

Therefore, under $\lambda_H = 1$ the best strategy of Self 0 is $\lambda_L = 0$ if $m \leq \frac{se_H q(\theta_H-\theta_L)}{q+\chi(r(1-q)-\epsilon)}$. Lemma

1 states that under $\lambda_H = 1$ and $\lambda_L = 0$, Self 1 chooses $e_L$. Hence, if the aim of Self 0 with $(\theta_L, \sigma_H)$ is to induce $e_H$, he can engage in a mixed strategy $\lambda_L = \rho$ where $\hat{\lambda}_H \leq \rho < 1$. This strategy is profitable if

$$\begin{aligned} \mathbb{E}_0\Big(U(e_L, \lambda_L = 1)\Big) \leq & \mathbb{E}_0\Big(U(e_H, \lambda_L = \rho)\Big) \Leftrightarrow \\ \theta_L(a+s)e_L \leq & \theta_L a e_H - c - m + \\ se_H\Bigg(\rho\theta_L + (1-\rho)\Big(\theta_H p_H(\lambda_H = 1, \lambda_L = \rho) & + \theta_L\Big(1 - p_H(\lambda_H = 1, \lambda_L = \rho)\Big)\Big)\Bigg), \end{aligned} \tag{19}$$

the second inequality in (19) can be rewritten as

$$\rho \leq 1 - \frac{\phi q}{1 - \phi\chi\left((1-q)r - \epsilon\right)}, \tag{20}$$

where $\phi := \frac{c - \theta_L(a+s) + m}{q(\theta_H - \theta_L)se_H}$. Suppose $\rho = \hat{\lambda}_H$. Replacing (11) into (20) yields

$$1 - \frac{(\theta_H(a+s) - c)q(r + (1-r)\lambda_H)}{(c - (a+s)\theta_L)\chi(1-q)r} \leq 1 - \frac{\phi q}{1 - \phi\chi\left((1-q)r - \epsilon\right)}, \tag{21}$$

which after some algebraic manipulations leads to

$$c \leq (a+s)\left(\frac{se_H q\theta_H + \chi\left((1-q)r - \epsilon\right)\left(m - (a+s)\theta_L\right)}{se_H q + (a+s)\chi\left((1-q)r - \epsilon\right)}\right) \tag{22}$$

Denote $\bar{m} := \frac{se_H q(\theta_H - \theta_L)}{q + \chi(r(1-q) - \epsilon)}$ and $\bar{c}$ as the right hand side of equation (22), then Self 0 with $(\theta_L, \sigma_H)$ chooses the mixed strategy $\lambda_L = \hat{\lambda}_L$ to which Self 1 reacts choosing $e_H$ if $m \leq \bar{m}$ and $c \leq \bar{c}$ hold.

Next, I show that it is unprofitable for Self 0 to deviate from $\lambda_L = \hat{\lambda}_H$ to opt for another mixed strategy such that $\hat{\lambda}_H < \rho < 1$ since

$$\begin{aligned} \mathbb{E}_0\Big(U(e_L, \lambda_L = \rho)\Big) \leq & \mathbb{E}_0\Big(U(e_H, \lambda_L = \hat{\lambda}_L)\Big) \Leftrightarrow \\ qse_H(\theta_H - \theta_L)\left(\frac{(1-\rho)}{q + (1-\rho)\chi\left((1-q)r - \epsilon\right)} - \right. & \left.\frac{(1-\hat{\lambda}_L)}{q + (1-\hat{\lambda}_L)\chi\left((1-q)r - \epsilon\right)}\right) \leq 0, \end{aligned} \tag{23}$$

where the second inequality in equation (23) is guaranteed by $\rho > \hat{\lambda}_L$. Therefore, to induce $e_H$, Self 0 with $(\theta_L, \sigma_H)$ sets the mixed strategy $\lambda_L^{**} = \hat{\lambda}_L$ if $m \leq \bar{m}$ and $c \leq \bar{c}$.

Furthermore, I show that under $\lambda_L^{**} = \hat{\lambda}_L$ Self 0 with $\theta_H$ sets $\lambda_H = 1$. Suppose that $\lambda_H = 0$ induces $e_H$. Then, Self 0 is better off with $\lambda_H = 1$ since he does not incur in the cost

$m$ and he obtains higher psychological benefits. Formally,

$$\begin{aligned}&\mathbb{E}_0\Big(U(\lambda_H=1,e_H)\Big)>\mathbb{E}_0\Big(U(\lambda_H=0,e_H)\Big)\\&\quad\theta_H(a+s)e_H-c>\theta_H(a+s)e_H-se_H(\theta_H-\theta_L)p(\lambda_H=0,\lambda_L=\hat{\lambda}_L)-m-c,\end{aligned}\tag{24}$$

where the second inequality in (24) holds since $-se_L(\theta_H-\theta_L)p(\lambda_H=0,\lambda_L=\hat{\lambda_L})-m<0$.

Suppose instead that $\lambda_H=0$ induces $e_L$. Then, Self 0 is again better off setting $\lambda_H=1$ since

$$\begin{aligned}&\mathbb{E}_0\Big(U(\lambda_H=1,e_H)\Big)>\mathbb{E}_0\Big(U(\lambda_H=0,e_H)\Big)\Leftrightarrow\\&\quad\theta_H(a+s)e_H-c>\theta_H(a+s)e_L-se_L(\theta_H-\theta_L)p(\lambda_H=0,\lambda_L=\hat{\lambda}_L)-m,\end{aligned}\tag{25}$$

where the second inequality in the equation above holds because $\theta_H(a+s)-c>0>-se_L(\theta_H-\theta_L)p(\lambda_H=\nu,\lambda_L=\hat{\lambda_L})-m$. Moreover, any mixed strategy $\lambda_H=\nu$ with $0<\nu<1$, yields lower psychological benefits than setting $\lambda_H=1$ and also implies self-deception costs $m>0$, making Self 0 worse off. Hence, Self 0 with $\theta_i=\theta_H$ is better off setting $\lambda_H^{**}=1$ whenever $\lambda_L^{**}=\hat{\lambda}_L$.

Finally, Selves 0 with $(\theta_L,\sigma_L)$ or with $(\theta_L,\sigma_L)$ are unable to engage in self-deception and set $\lambda_L^{**}=1$ and $\lambda_H^{**}=1$. These strategies yield $e_L$ and $e_H$, respectively.

■

**Corollary 1**

*Proof.* Under $\epsilon=0$, $\bar{m}=\frac{se_Hq(\theta_H-\theta_L)}{q+\chi(r(1-q))}$. Since $\theta_H>\theta_L$ then $\bar{m}>0$, which ensures the existence of an interval in which $m<\bar{m}$. Moreover, under $\epsilon=0$,

$$\bar{c}=(a+s)\left(\frac{se_Hq\theta_H+\chi\left((1-q)r\right)(m-(a+s)\theta_L)}{se_Hq+(a+s)\chi\left((1-q)r\right)}\right).\tag{26}$$

From the above equation it is evident that a sufficient condition for $\bar{c}>0$ is $m>(a+s)\theta_L$.

I next show that these two conditions, which imply $\frac{se_Hq(\theta_H-\theta_L)}{q+\chi(r(1-q))}>m>(a+s)\theta_L$, are compatible. It suffices to show that the following inequality is feasible

$$\frac{se_Hq(\theta_H-\theta_L)}{q+\chi\left(r(1-q)\right)}>(a+s)\theta_L,\tag{27}$$

which holds if $\frac{a+s}{s}<\left(\frac{\theta_H}{\theta_L}-1\right)e_Hq$. Since $\theta_H>\theta_L$ a condition that holds for a sufficiently large value of $s$, or $e_H$.

**Corollary 2**

*Proof.* The partial derivative of $\epsilon$ with respect to $\bar{m}$ is:

$$\frac{\partial \bar{m}}{\partial \epsilon} = \frac{\chi s e_L q(\theta_H - \theta_L)}{\left(q + \chi\left(r(1-q) - \epsilon\right)\right)^2} \tag{28}$$

Since $\theta_H > \theta_L$, $\frac{\partial \bar{m}}{\partial \epsilon} > 0$. Therefore, the condition $m < \bar{m}$ becomes less stringent as $\epsilon$ increases.

Moreover, the partial derivative of $\epsilon$ with respect to $\bar{c}$ is:

$$\frac{\partial \bar{c}}{\partial \epsilon} = \frac{\chi(a+s)}{s e_H q + (a+s)\chi((1-q)r - \epsilon)}\Big((a+s)\theta_L + \bar{c} - m\Big) \tag{29}$$

Since $r(1-q) \geq \epsilon$, a necessary and sufficient condition for $\frac{\partial \bar{c}}{\partial \epsilon} < 0$ is $m > (a+s)\theta_L + \bar{c}$

Let $\frac{\partial \bar{c}}{\partial \epsilon} < 0$. Since $r(1-q) \geq \epsilon$, equation(29) shows that a necessary requirement for attaining that sign is

$$m > (a+s)\theta_L + \bar{c}. \tag{30}$$

Replacing $\bar{c}$ from Proposition 1 into (30) leads to the condition

$$m > (a+s)(\theta_H + \theta_L). \tag{31}$$

Notice that for given $c$ and $m$ $\max\left(U_1(e_H, \lambda) - U_1(e_L, 1)\right) = (a+s)\theta_i - m - c$. Equation (31) implies that $e_H$ cannot be attained, contradicting the existence of Proposition 1. Then, it must be that $\frac{\partial \bar{c}}{\partial \epsilon} > 0$.

■

**Proposition 2**

*Proof.* Suppose that Self 0 with $\theta_L$ chooses $\lambda_L = 1$. Lemma 1 states that after receiving $\tilde{\theta}_H$, Self 1 chooses $e_H$ since $\lambda_L = 1 \geq \hat{\lambda}_L$. Instead, Self 1 chooses either $e_H$ or $e_L$ after receiving $\tilde{\theta}_L$. Assume that $\lambda_L = 0$ induces Self 1 to choose $e_H$. Then, Self 0 with $(\theta_H, \sigma_L)$ is better off choosing $\lambda_H = 1$ as long as:

$$\begin{aligned} \mathbb{E}_0\Big(U(\lambda_H = 1, e_H)\Big) \geq & \mathbb{E}_0\Big(U(\lambda_H = 0, e_H)\Big) \Leftrightarrow \\ \theta_H(a+s)e_H - c \geq & \theta_H(a+s)e_H - s e_H(\theta_H - \theta_L)p_L(\lambda_H = 0, \lambda_L = 1) - m - c, \end{aligned} \tag{32}$$

where the second inequality above holds since $-s e_L(\theta_H - \theta_L)p_L(\lambda_H = 0, \lambda_L = 1) - m < 0$.

Alternatively, let $\lambda_H = 0$ yield $e_L$. Then, Self 0 with $(\theta_H, \sigma_L)$ is better off choosing $\lambda_H = 1$

as long as:

$$
\begin{aligned}
&\mathbb{E}_0\Big(U(\lambda_H = 1, e_H)\Big) \geq \mathbb{E}_0\Big(U(\lambda_H = 0, e_H)\Big) \Leftrightarrow \\
&\quad \theta_H(a+s)e_H - c \geq \theta_H(a+s)e_L - se_L(\theta_H - \theta_L)p_L(\lambda_H = 0, \lambda_L = 1) - m,
\end{aligned}
\tag{33}
$$

where the second inequality above holds for all the feasible values that the parameters of the model can attain since $\theta_H(a+s) - c > 0 > -se_L(\theta_H - \theta_L)p_L(\lambda_H = 0, \lambda_L = 1) - m$. Hence, Self 0's best response is to set $\lambda_H^{**} = 1$ if $\lambda_L = 1$.

Suppose now that $\lambda_H = 1$. According to Lemma 1, Self 1 chooses $e_L$ upon receiving $\tilde{\theta}_L$ since $\lambda_H = 1 \geq \hat{\lambda}_H$. Instead, if Self 1 receives $\tilde{\theta}_H$ he can either choose $e_H$ or $e_L$. The strategy of Self 0 with the tuple ability-status $(\theta_L, \sigma_H)$ depends on the magnitude of $m$. To see how, suppose that $\lambda_L = 0$ induces $e_L$, $\lambda_L = 1$ makes Self 0 better off as long as:

$$
\begin{aligned}
&\mathbb{E}_0\Big(U(\lambda_L = 1, e_L)\Big) > E_0\Big(U(\lambda_L = 0, e_L)\Big) \Leftrightarrow \\
&\quad \theta_L(a+s)e_L > \theta_L(a+s)e_L + se_L(\theta_H - \theta_L)p_H(\lambda_H = 1, \lambda_L = 0) - m,
\end{aligned}
\tag{34}
$$

which holds whenever $m > \frac{se_L q(\theta_H - \theta_L)}{q + \chi((1-q)r - \epsilon)}$

Alternatively, when $\lambda_L = 0$ induces $e_H$, Self 0 is better off setting $\lambda_L = 1$ whenever:

$$
\begin{aligned}
&\mathbb{E}_0\Big(U(\lambda_L = 1, e_L)\Big) > \mathbb{E}_0\Big(U(\lambda_L = 0, e_H)\Big) \Leftrightarrow \\
&\quad \theta_L(a+s)e_L > \theta_L(a+s)e_H + se_H(\theta_H - \theta_L)p_H(\lambda_H = 1, \lambda_L = 0) - m - c,
\end{aligned}
\tag{35}
$$

which holds as long as $m > \frac{se_H q(\theta_H - \theta_L)}{(q + \chi(r(1-q)) - \epsilon)}$. Therefore, Self 0's best strategy is to set $\lambda_L^{**} = 1$ if $\lambda_H = 1$ and $m > \frac{se_H q(\theta_H - \theta_L)}{q + \chi(r(1-q) - \epsilon)}$.

Finally, note that Lemma 1 states that when $\lambda_L^{**} = 1$ and $\lambda_H^{**} = 1$, Self 1 reacts to $\tilde{\theta}_L$ with $e_L$ and to $\tilde{\theta}_H$, with $e_H$.

■

## Appendix B: Additional equilibria of the status and self-deception game

In this appendix I present the remaining equilibria of the social status and self-deception game. These equilibria are relegated to an appendix for two reasons. The first is to keep the paper short and coherent; presenting the whole set of equilibria could deviate the reader's attention from the main message of this study which that status can have an influence on performance by means of beliefs. This message is adequately conveyed by Proposition 1. The second reason is that the equilibria presented in the main body of the paper are the empirically relevant equilibria. While Proposition 1 yields the message that status can affect beliefs and performance, Proposition 2 represents the opposing situation in which status is economically irrelevant. Instead, the equilibria described below yield that status affects beliefs and performance in unexpected ways.

I start describing an equilibrium in which individuals exert low effort regardless of their social status or ability. This equilibrium is a pooling equilibrium at low levels of effort and it presents a situation in which low ability individuals engage in self-deception whenever they can, making favorable signals non-credible.

**Definition 3.** *A "low pooling" equilibrium of the status and self-deception game is characterized by the tuple* $(\lambda_\ell^{**}, e_\ell^{**})$*, where*

$$e_\ell^{**} = e_L$$

*and*

$$\lambda_\ell^{**} = \begin{cases} 0 & \text{if } (\theta_L, \sigma_H), \\ 1 & \text{if } \theta_H \text{ or } (\theta_L, \sigma_L). \end{cases}$$

**Proposition 3.** *The equilibrium* $(\lambda_\ell^{**}, e_\ell^{**})$ *is sustained if* $s > 0$*,* $c > \bar{c}$*,* $m \leq \bar{m}$*, and* $m > a(\theta_H - \theta_L) + s(\theta_H - \theta_L)\frac{((1-q)(1-r)+\epsilon)e_L + \chi((1-q)r-\epsilon)}{(1-q)(1-r)+\epsilon+\chi((1-q)r-\epsilon)}$.

*Proof.* Suppose that $\lambda_H = 1$. According to Lemma 1, Self 1 chooses $e_L$ when receiving $\tilde{\theta}_L$ since $\lambda_H = 1 \geq \hat{\lambda}_H$. Instead, after receiving $\tilde{\theta}_H$, Self 1 chooses either $e_H$ or $e_L$. Given these possible reactions, Self 0's strategy depends on the magnitude of $m$. In particular, when $\lambda_L = 0$ induces $e_L$ and $m \leq \frac{se_L q(\theta_H - \theta_L)}{q+\chi(r(1-q)-\epsilon)}$, Self 0 is better off setting $\lambda_L = 0$, because for those values of m, $\mathbb{E}_0\Big(U(\lambda_L = 1, e_L)\Big) = \theta_L(a+s)e_L \leq \theta_L(a+s)e_L + se_L(\theta_H - \theta_L)p_H(\lambda_H = 1, \lambda_L = 0) - m = \mathbb{E}_0\Big(U(\lambda_L = 0, e_L)\Big)$. Instead, if $\lambda_L = 0$ induces $e_H$, Self 0 is better off setting $\lambda_L = 0$ as long as $m \leq \frac{se_H q(\theta_H - \theta_L)}{q+\chi(r(1-q)-\epsilon)}$ since $\mathbb{E}_0\Big(U(\lambda_L = 1, e_L)\Big) = \theta_L(a+s)e_L \leq$

$\theta_L(a+s)e_H + se_H(\theta_H-\theta_L)p_H(\lambda_H=1,\lambda_L=0) - m - c = \mathbb{E}\Big(U(\lambda_L=0,e_H)\Big)$ requires that $m$ attains, at least, $m = \frac{se_H q(\theta_H-\theta_L)}{q+\chi((1-q)-\epsilon)}$.

Additionally, according to Proposition 1, Self 0 prefers to set $\lambda_L = 0$ to $\lambda_L = \hat{\lambda}_L$ if $c > \bar{c}$, since $\mathbb{E}_0\Big(U(e_L,\lambda_L=0)\Big) > \mathbb{E}_0\Big(U(e_H,\lambda_L=\hat{\lambda}_H)\Big)$. Therefore, Self 0's best strategy is to set $\lambda_L = 0$ if $\lambda_H = 1$ and $m \leq \frac{se_L q(\theta_H-\theta_L)}{q+\chi(r(1-q)-\epsilon)}$.

Next, suppose that $\lambda_L = 0$. According to Lemma 1, Self 1 chooses $e_L$ after receiving $\tilde{\theta}_H$ since $\lambda_L = 0 < \hat{\lambda}_L$. Instead, if $\tilde{\theta}_L$ is received, Self 1 chooses either $e_H$ or $e_L$. Self 0 is better off setting $\lambda_H = 1$ if $\lambda_H = 0$ induces $e_L$, since $\mathbb{E}_0\Big(U(\lambda_H=1,e_L)\Big) = \theta_H(a+s)e_L \geq \theta_H(a+s)e_L - se_L(\theta_H-\theta_L)p_L(\lambda_H=0,\lambda_L=0) - m = \mathbb{E}_0\Big(U(\lambda_H=0,e_L)\Big)$ always holds due to $-se_L(\theta_H-\theta_L)p_L(\lambda_H=0,\lambda_L=0) - m < 0$. In contrast, when $\lambda_H = 0$ induces $e_H$, Self 0 is better off with $\lambda_H = 1$ rather than with $\lambda_H = 0$ whenever $m > a(\theta_H-\theta_L) + s(\theta_H-\theta_L)\frac{((1-q)(1-r)+\epsilon)e_L+\chi((1-q)r-\epsilon)}{(1-q)(1-r)+\epsilon+\chi((1-q)r-\epsilon)}$. This is because $\mathbb{E}_0\Big(U(\lambda_H=1,e_L)\Big) = \theta_H(a+s)e_L > \theta_H(a+s)e_H - se_H(\theta_H-\theta_L)p_L(\lambda_H=0,\lambda_L=0) - m - c = \mathbb{E}_0\Big(U(\lambda_H=0,e_H)\Big)$ holds for such values of $m$. Therefore, when $\lambda_L = 0$ and the cost of self-deception is large, Self 0 with $(\theta_H, s_L)$ sets $\lambda_H^{**} = 1$.

Lemma 1 states that when $\lambda_L^{**} = 0$ and $\lambda_H^{**} = 1$, favorable signals $\tilde{\theta}_H$ are best responded with $e_L$. Finally, individuals with $(\theta_L,\sigma_L)$ and $(\theta_H,\sigma_H)$ set $\lambda_L^{**} = 1$ and $\lambda_H^{**} = 1$, respectively, to which Self 1 reacts with $e_L$.

■

Proposition 3 presents an equilibrium in which low ability individuals engage in a pure strategy of self-deception $\lambda_L = 0$. In contrast to Proposition 1, engaging in a mixed strategy $\lambda_L = \hat{\lambda}_H$ to make Self 1 indifferent is no longer profitable due to the high costs faced by the individual when the high effort $e_H$ is chosen. However, self-deception is profitable inasmuch as it yields psychological utility and the costs associated to it are moderate for the low types, i.e. $m < \bar{m}$. For high ability individuals is best strategy to be truthful because they could generate $e_L$ without the need of engaging in self-deception and because the costs associated to high effort are too elevated. Finally, Self 1 reacts to favorable signals by setting low effort levels.

Another equilibrium arising from the game is a separating equilibrium in social status. However, in this equilibrium effort becomes lower with higher status.

**Definition 4.** *A separating equilibrium of the status and self-deception game is characterized by the tuple* $(\lambda_m^{**}, e_m^{**})$*, where*

$$e_m^{**} = \begin{cases} e_L & \textit{if } (\theta_L,\sigma_H) \textit{ and } (\theta_H,\sigma_H), \\ e_H & \textit{if } (\theta_L,\sigma_L) \textit{ and } (\theta_H,\sigma_L). \end{cases}$$

*and*

$$\lambda_m^{**} = \begin{cases} 0 & \textit{if } (\theta_L, \sigma_H) \textit{ and } (\theta_H, \sigma_L), \\ 1 & \textit{if } (\theta_L, \sigma_L) \textit{ and } (\theta_H, \sigma_H). \end{cases}$$

**Proposition 4.** *The equilibrium* $(\lambda_m^{**}, e_m^{**})$ *is sustained if* $s > 0$*,* $c > \bar{c}$*,* $m < \bar{m}$*, and* $m \leq a(\theta_H - \theta_L) + s(\theta_H - \theta_L)\frac{((1-q)(1-r)+\epsilon)e_L + \chi((1-q)r-\epsilon)}{(1-q)(1-r)+\epsilon+\chi((1-q)r-\epsilon)}$*.*

*Proof.* Let $\lambda_H = 0$. According to Lemma 1, Self 1 chooses $e_H$ after receiving $\tilde{\theta}_L$ since $\lambda_H = 0 < \hat{\lambda}_H$. Instead, if $\tilde{\theta}_H$ is received, Self 1 chooses either $e_H$ or $e_L$. Self 0's strategy with $(\theta_L, \sigma_H)$ depends on $m$. Let $\lambda_L = 0$ generate $e_L$, then $\mathbb{E}_0\Big(U(\lambda_L = 0, e_L\Big) = (a + s)\theta_L e_L + s e_L(\theta_H - \theta_L)p_H(\lambda_H = 0, \lambda_L = 0) - m \geq (a+s)\theta_L e_H - c = \mathbb{E}_0\Big(U(\lambda_L = 1, e_H\Big)$ whenever $m \leq \frac{(qr+\epsilon)s e_L(\theta_H - \theta_L)}{qr+\epsilon+\chi((1-q)r-\epsilon)}$. Next, let $\lambda_L = 0$ induce $e_H$. Self 0's strategy again depends on the values that $m$ attains. Strategy $\lambda_L = 0$ is profitable if $m \leq \frac{(qr+\epsilon)s e_H(\theta_H - \theta_L)}{qr+\epsilon+\chi((1-q)r-\epsilon)}$. That $m$ attains such lower bound guarantees that $\mathbb{E}_0\Big(U(\lambda_L = 1, e_H)\Big) = \theta_L(a+s)e_H - c \leq \theta_L(a+s)e_H + s e_H(\theta_H - \theta_L)p_H(\lambda_H = 0, \lambda_L = 0) - m - c = \mathbb{E}_0\Big(U(\lambda_L = 0, e_L)\Big)$.

Let $\lambda_L = 0$. According to Lemma 1, Self 1 chooses $e_L$ after receiving $\tilde{\theta}_H$ since $\lambda_L = 0 < \hat{\lambda}_L$. Instead, if $\tilde{\theta}_L$ is received, Self 1 chooses either $e_H$ or $e_L$. Suppose that $\lambda_H = 0$ induces $e_L$. Then, self 0 is better off setting $\lambda_H = 1$ since the inequality $\mathbb{E}_0\Big(U(\lambda_H = 1, e_L)\Big) = \theta_H(a+s)e_L \geq \theta_H(a+s)e_L - s e_L(\theta_H - \theta_L)p_L(\lambda_H = 0, \lambda_L = 0) - m = \mathbb{E}_0\Big(U(\lambda_H = 0, e_L)\Big)$ holds for all the possible values that the parameters of the model can attain since $-s e_L(\theta_H - \theta_L)p_L(\lambda_H = 0, \lambda_L = 0) - m < 0 < \theta_H(a+s)$. Instead,Self 0 is better off setting $\lambda_H = 0$ rather than $\lambda_H = 1$ if $\lambda_H = 0$ induces $e_H$ and $m \leq a(\theta_H - \theta_L) + s(\theta_H - \theta_L)\frac{((1-q)(1-r)+\epsilon)e_L + \chi((1-q)r-\epsilon)}{(1-q)(1-r)+\epsilon+\chi((1-q)r-\epsilon)}$. Because for such values of $m$ the inequality $\mathbb{E}_0\Big(U(\lambda_H = 1, e_L)\Big) = \theta_H(a+s)e_L \leq \theta_H(a+s)e_H - s e_H(\theta_H - \theta_L)p_L(\lambda_H = 0, \lambda_L = 0) - m - c = \mathbb{E}_0\Big(U(\lambda_H = 0, e_H)\Big)$ holds. Thus, when $\lambda_L = 0$ Self 0 with $\theta_i = \theta_H$ is better off sending $\lambda_H = 1$ as long as the cost of self-deception is not larger than the upper bound $m \leq a(\theta_H - \theta_L) + s(\theta_H - \theta_L)\frac{((1-q)(1-r)+\epsilon)e_L + \chi((1-q)r-\epsilon)}{(1-q)(1-r)+\epsilon+\chi((1-q)r-\epsilon)}$.

Lemma 1, shows that for $\lambda_L^{**} = 0$ and $\lambda_H^{**} = 1$, Self 1 reaction to receiving $\tilde{\theta}_H$ is setting $e_L$. Also, upon receiving $\tilde{\theta}_L$ Self 1's best-response is to set $e_H$. Finally, individuals with $(\theta_L, \sigma_L)$ and $(\theta_H, \sigma_H)$ set $\lambda_L^{**} = 1$ and $\lambda_H^{**} = 1$, to which self 1 reacts with $e_H$ and $e_L$, respectively.

■

Proposition 4 presents an equilibrium in which both high and low ability individuals engage in pure strategies of self-deception whenever possible. Such behavior makes favorable and unfavorable signals not credible. According to Lemma 1, this propensity to engage in self-deception from both types entails that Self 1's reaction to favorable signals is to exert low effort, $e_L$ and also that the reaction to unfavorable signals is to set high effort, $e_H$.

This equilibrium emerges because low ability individuals have no incentive to be truthful: they derive utility gains from holding high beliefs about their ability. Similarly, high ability individuals engage in a strategy of self-deception because self-deception is not too costly and because being truthful generates lower effort levels, which leads to lower utility levels.

## Appendix C: Description of control variables from survey data analyses

Table 8: Variables description

| Variable | Sweep | Type | Description |
|---|---|---|---|
| *Demographic* | | | |
| Mother's age birth | 1970 | Continuous | Age of mother at first birth |
| Father's age | 1970 | Continuous | Age of father at present marriage |
| Female | 1970 | Binaty | Whether individual is female |
| Family Income | 1980 | Ordinal | Gross weekly family income |
| *Cognitive ability* | | | |
| Number of O levels | 1986 | Continuous | Number of achieved O-levels |
| EPVT | 1975 | Continuous | English Picture Vocabulary Test |
| MATHS | 1980 | Continuous | Friendly Maths Test Score |
| READ | 1980 | Continuous | Edinburgh Reading Test Score |
| Copy | 1975 | Continuous | Score of Copying Test |
| BAST | 1980 | Continuous | British Ability Scale Test |
| *Non-cognitive traits* | | | |
| Anti-social | 1975 | Continuous | Disobedient, destructive, aggressive, restless and tantrum. |
| Neurotic | 1975 | Continuous | miserable, worried, fearful, fussy and complains of aches and pains. |
| Anxiety | 1980 | Continuous | 9 items on anxiety. |
| Hyperactivity | 1980 | Continuous | 6 items, includes the items squirmy, excitable, twitches, hums and taps. |
| Application | 1980 | Continuous | Child's concentration and perseverance and ability to understand and complete complex tasks. |
| Clumsiness | 1980 | Continuous | items on bumping into things, and the use of small objects such as scissors. |
| Extroversion | 1980 | Continuous | items concerning talkativeness and an explicit question about extroversion. |
| CAROLOC | 1980 | Continuous | Score for locus of control (Gammage, 1975) |
| LAWSEQ | 1980 | Continuous | Score for self-confidence (Lawrence, 1973, 1978) |

Variables description (Continued)

| Variable | Sweep | Type | Description |
|---|---|---|---|
| *Self-reported skills* | | | |
| Communicator | 1986 | Binary | Whether individual is a good communicator |
| Handy | 1986 | Binary | Whether individual is good with hands |
| Thinker | 1986 | Binary | Whether individual is a clear thinker |
| Worker | 1986 | Binary | Whether individual is a hard worker |
| Tidy | 1986 | Binary | Whether individual is clean and tidy |
| Reliable | 1986 | Binary | Whether individual is reliable |
| Time | 1986 | Binary | Whether individual is a good time-keeper |
| Responsible | 1986 | Binary | Whether individual can take responsibility |
| *Attitudes toward school* | | | |
| Waste | 1986 | Categorical | School is largely a waste of time |
| Quiet | 1986 | Categorical | Quiet in classroom and get on with work |
| Bore | 1986 | Categorical | Thinks homework is a bore |
| Mind | 1986 | Categorical | Find it difficult to keep mind on work |
| Seriously | 1986 | Categorical | Never take work seriously |
| Do not like school | 1986 | Categorical | Do not like school |
| No plans | 1986 | Categorical | Plans pointless, take things as they come |
| Help teacher | 1986 | Categorical | Always willing to help the teacher |
| *Health variables* | | | |
| Missed school | 1986 | Binary | Missed school past 2 yrs due to health |
| Days Missed | 1986 | Continuous | Number of school days lost |
| Depressed | 1986 | Binary | Whether individual depressed or anxious |
| Days depressed | 1986 | Continuous | Days the individual is depressed |
| Medical help | 1986 | Binary | Visited the doctor when depressed |
| *Neighborhood variables* | | | |
| Neighbors | 1986 | Categorical | What are people like in your neighborhood? |
| Walk alone | 1986 | Categorical | Ever walk alone in your area after dark? |
| Safe walk alone | 1986 | Categorical | if walk alone, how do you feel? |
| *Expectations* | | | |
| Parents Expectations | 1986 | Binary | Advised by parents to leave school early. |
| Teachers Expectations | 1986 | Binary | Advised by teachers to leave school early. |
| Own expectations | 1986 | Binary | Always taken early leaving for granted |

Table 9: Descriptive statistics of variables

| **Variable** | **N** | **Mean** | **Median** | **St. Dev.** |
|---|---|---|---|---|
| *Demographics* | | | | |
| Mothers' age birth | 17,196 | 21.94 | 22 | 4.727 |
| Father's age | 17,196 | 13.551 | 21 | 14.939 |
| Female | 8,874 | 0.526 | 1 | 0.499 |
| Family Income | 14,875 | 2.612 | 3 | 1.993 |
| *Cognitive ability* | | | | |
| No. of O-levels | 23,327 | 1.263 | 0 | 2.449 |
| EPVT | 13,1357 | 34.280 | 36 | 16.256 |
| MATHS | 14,875 | 34.154 | 40 | 21.521 |
| READ | 14,875 | -0.342 | -0.76 | 0.940 |
| Copy | 13,135 | 4.663 | 5 | 2.089 |
| BAST | 11,368 | 62.662 | 63 | 12.295 |
| *Non-cognitive traits* | | | | |
| Anti-social | 13,135 | 9.228 | 9 | 3.825 |
| Anxiety | 12,751 | 135.374 | 124 | 76.910 |
| Hyperactivity | 12,751 | 69.969 | 58 | 50.655 |
| Neurotic | 13,135 | 8.044 | 8 | 3.392 |
| Application | 12,751 | 312.581 | 311 | 59.116 |
| Clumsiness | 12,751 | 197.218 | 195 | 47.137 |
| Extroversion | 12,751 | 145.608 | 147 | 27.123 |
| CAROLOC | 23,327 | 3.675 | 3 | 3.942 |
| LAWSEQ | 23,327 | 8.247 | 8 | 8.237 |
| *Self-reported skills* | | | | |
| Communicator | 11,617 | 0.459 | 0 | 0.36 |
| Handy | 11,617 | 0.495 | 0 | 0.387 |
| Thinker | 11,617 | 0.494 | 0 | 0.387 |
| Worker | 11,617 | 0.299 | 0 | 0.221 |
| Tidy | 11,617 | 0.432 | 0 | 0.340 |
| Reliable | 11,617 | 0.295 | 0 | 0.217 |
| Responsible | 11,617 | 0.363 | 0 | 0.644 |

Descriptive statistics of variables (continued)

| Variable | N | Mean | Median | St. Dev. |
|---|---|---|---|---|
| *Attitudes toward school* | | | | |
| Waste | 11,617 | 0.952 | 2 | 1.879 |
| Quiet | 11,617 | 0 .604 | 1 | 1.564 |
| Bore | 11,617 | 0.490 | 1 | 1.480 |
| Mind | 11,617 | 0.679 | 1 | 1.650 |
| Seriously | 11,617 | 0.910 | 2 | 1.848 |
| No Plans | 11,617 | 0.907 | 1 | 1.854 |
| Help teacher | 11,617 | 0.636 | 1 | 1.603 |
| *health variables* | | | | |
| Missed school | 11,617 | 0.544 | 1 | 0.438 |
| Days missed | 11,617 | 2.014 | 0 | 12.055 |
| Depressed | 11,617 | 0.590 | 0 | 0.426 |
| Days depressed (>0) | 11,617 | 3.308 | 4 | 0.871 |
| Medical help (>0) | 11,617 | 1.925 | 1 | 0.262 |
| *Neighborhood* | | | | |
| Neighbors | 11,617 | .996 | 1 | 1.816 |
| Walk alone | 11,617 | .996 | 1 | 1.209 |
| Safe walk alone | 11,617 | 0 .837 | 1 | 1.689 |
| *Expectations* | | | | |
| Parents Expectations (>-1) | 11,617 | 0.918 | 1 | 0.274 |
| Teacher Expectations(>-1) | 11,617 | 0.96 | 1 | 0.171 |
| Own expectations(>-1) | 11,617 | 0.903 | 1 | 0.087 |

# Appendix D: Additional Analyses

## D.1 Cohort data with discrete variables

This Appendix presents the analysis of the cohort data when the most relevant variables are discrete. The variables *"High Status Parents"*, *"High Aspirations"*, and *"High Status"*, indicate either occupations that require non-vocational tertiary education levels such as university, or the aspiration to achieve these education levels. The descriptive statistics of these variables are presented in Table 10

Table 10: Descriptive statistics of discrete variables

| **Variable** | Mean | St. Dev. | Median | Max. | Min. |
|---|---|---|---|---|---|
| High Aspirations | 0.417 | 0.417 | 0 | 1 | 0 |
| High Status Parents | 0.182 | 0.386 | 0 | 1 | 0 |
| High Status Adult | 0.412 | 0.492 | 0 | 1 | 0 |

Note: This table presents the averages, standard deviations, medians, the maximum value and the minimum value of the main variables used in the survey analysis.

The estimates of logistic regressions featuring the discrete versions of the variables of interest are presented in Table 11. Columns 1 and 2 show that belonging to a high status is associated a 5.4% higher chance to have high educational aspirations and this increase is significant at the 1% level for all specifications. Hence, I find empirical evidence supporting the claim that higher status at birth is related with a higher educational aspirations and that this relationship is robust to controlling for variables that reflect the individuals' aptitudes and circumstances.

The estimates presented in Columns 3 and 4 in Table 11 suggest that belonging to a high status is associated a 5.4% higher chance to have high educational aspirations and this increase is significant at the 1% level for all specifications. Hence, I find empirical evidence supporting the claim that higher status at birth is related with a higher educational aspirations and that this relationship is robust to controlling for variables that reflect the individuals' aptitudes and circumstances.

## D.2 Regressions Experiment 1

To account for factors other than the treatment assignment that might influence subjects' performance, I perform regression analyses that seek to evaluate the treatment effects while

Table 11: Determinants of High Achieved Status and High Aspirations

| | (1) High Aspirations | (2) High Aspirations | (3) High Status Adult | (4) High Status Adult |
|---|---|---|---|---|
| High Aspirations* High Status Parents | | | 2.097*** | 0.884*** |
| | | | (0.162) | (0.203) |
| High Aspirations | | | 1.361*** | 0.836*** |
| | | | (0.114) | (0.146) |
| High Status Parents | 0.892*** | 0.311** | 0.134 | -0.142 |
| | (0.097) | (0.122) | (0.163) | (0.199) |
| Mother's age at birth | 0.005 | -0.007 | 0.008 | 0.004 |
| | (0.005) | (0.006) | (0.006) | (0.007) |
| Female | 0.492*** | 0.428*** | -0.568*** | -0.550*** |
| | (0.083) | (0.127) | (0.099) | (0.138) |
| Family income 1980 | | -0.142*** | | -0.113*** |
| | | (0.030) | | (0.032) |
| Constant | 1.975*** | 1.614* | -1.026*** | -2.575** |
| | (0.307) | (0.960) | (0.384) | (1.074) |
| Cognitive skills | NO | YES | NO | YES |
| Non-cognitive traits | NO | YES | NO | YES |
| Self-reported skills | NO | YES | NO | YES |
| Neighborhood | NO | YES | NO | YES |
| Expectations | NO | YES | NO | YES |
| Attitudes toward school | NO | YES | NO | YES |
| Health | NO | YES | NO | YES |
| Country Fixed effects | YES | YES | YES | YES |
| Region Fixed effects | YES | YES | YES | YES |
| Log-Likelihood | -1728.591 | -1095.687 | -1271.931 | -936.744 |
| Observations | 2661 | 2178 | 2,090 | 1,719 |

Note: Columns 1, 2, and 3 presents the estimates of the ordered logistic regression of the model $HighAspirations_i = \beta_0 + \beta_1 HighStatusParents_i + Controls'\Gamma + \epsilon_i$ with $\epsilon_i \sim logistic$. Columns 4, 5, and 6 presents the estimates of the ordered logistic regression of the model and $HighStatusAdult_i = \beta_0 + \beta_1 HighStatusParents * HighAspirations + \beta_2 HighAspirations + \beta_3 HighStatusParents_i + Controls'\Gamma + \epsilon_i$ with $\epsilon_i \sim logistic$. "High Status Adult" is an binary variable that takes a value of one if the individual's occupation at the age of 37 required either university or technical education. "High Aspirations" is a binary variable that takes a value of one when the educational aspirations of the individual at adolescence indicated either university or technical education. "High Status Parents" is a binary variable that takes a value of one when the parents' education was either university or technical education. The description of the control variables is presented in Appendix C. Clustered standard errors at the individual level in parentheses. *** denotes significance at the 0.01 level, ** denotes significance at the 0.05 level, * denotes significance at the 0.1 level.

controlling for a set of relevant variables.[26] Columns 1 and 2 in Table 12 present the estimates of negative binomial regressions of performance in Part 2 on subjects' ability, a treatment dummy, the interaction between these two variables, and control variables.[27]

The regression estimates confirm the aforementioned results. Specifically, among low ability subjects, those assigned high status attain higher average performance ($\chi^2(1) = 9.20$, p = 0.002). Moreover, the estimate associated to High Status shows that the treatments did not generate differences in average performance among high ability subjects. In addition, the estimate associated toLow ability corroborates that subjects with low ability assigned to the low status exhibit lower average performance than high ability subjects assigned to the same treatment. Finally, and in contrast to the analysis of the data based on pairwise testing, I find that high ability subjects assigned to High Status were outperformed by low ability subjects assigned to the same treatment ($\chi^2(1)$= 5.20, p = 0.022).

Columns 3 and 4 in Table 4 present the estimates of negative binomial regressions of beliefs on treatment dummies and, for some specifications, control variables. The resulting estimates confirm the aforementioned findings. First, low ability subjects exhibit lower average beliefs than high ability subjects when both groups are assigned to Low Status (p = 0.025). Second, average beliefs do not differ among subjects of low and high ability when both groups are assigned to the High Status treatment ($\chi^2(1)$=0.60, p=0.437). Third, and more importantly, subjects with a low status treatment exhibit higher average beliefs when assigned to High Status ($\chi^2(1)$=3.04, p= 0.04). Finally, I find no empirical evidence of a difference in beliefs between high ability subjects assigned to different treatments (p = 0.906).

### D.3 Heterogeneous treatment effects

To better understand the effect of the treatments on performance, I abandon the binary classification of ability, e.g. high and low, and replace it by the subjects' performance on the first part of the experiment. The latter measure can be interpreted as a "continuous" measure of ability, and it allows for a richer quantification of the treatment effect. Table 13 presents the estimates of negative binomial regressions of performance on the second part of the experiment on performance on the first part of the experiment, a treatment dummy and, for some specifications, relevant controls.

The resulting estimates suggest that the treatments generate a drastic change in the

[26]The matrix of control variables included in the regression contains the variables gender, number of participants in a given session, score in the locus of control questionnaire, subjects' beliefs about performance in Part 1 of the experiment, subjects' perception about the fairness of the treatment assignment, and whether subjects performed the task in the past.

[27]The performance data exhibit a variance, 69.30, that is larger than its mean, 23.434 A standard count regression model does not account for a data process with these characteristics. To account for this over-dispersion, I use a negative binomial model.

Table 12: Treatment Effects in Experiment 1

| | (1) Performance | (2) Performance | (3) Beliefs | (4) Beliefs |
|---|---|---|---|---|
| Low Ability | -5.095*** | -6.123*** | -4.702* | -5.353** |
| | (1.862) | (1.984) | (2.502) | (2.358) |
| High Status | -2.198 | -2.260 | -0.438 | -0.221 |
| | (1.809) | (1.784) | (2.089) | (1.873) |
| Low Ability*High Status | 6.930** | 7.146** | 3.257 | 3.248 |
| | (2.817) | (2.787) | (3.287) | (3.086) |
| Session size | | -0.601* | | 0.040 |
| | | (0.354) | | (0.429) |
| Belief practice round | | -0.018 | | 0.335 |
| | | (0.195) | | (0.271) |
| Medal fair | | -2.605 | | -2.546 |
| | | (1.690) | | (1.739) |
| Locus of Control | | 0.725* | | 0.274 |
| | | (0.420) | | (0.525) |
| Female | | 3.594** | | -0.630 |
| | | (1.644) | | (1.970) |
| Task performed before | | 6.632*** | | 6.258** |
| | | (2.032) | | (2.781) |
| ln($\delta$) | 0.658*** | 0.4190*** | 0.568*** | 0.424*** |
| | (0.265) | (0.297) | (0.205) | (0.232) |
| N | 133 | 133 | 133 | 133 |
| Log-likelihood | -463.927 | -454.243 | -480.819 | -475.162 |

Note: This table presents marginal effects of negative binomial regressions of the model $y_i = \beta_0 + \beta_1 LowAbility + \beta_2 HighStatus + \beta_3 LowAbility * HighStatus + Controls'\Gamma + \epsilon_i$, with $\epsilon \sim poisson(\lambda)$. $y_i$ is "Performance" in column (1) and column (2) and "Beliefs" in column (3) and column (4). "Performance" is the number of correctly solved matrices in the second part of Experiment 1, "Beliefs' is the number of correctly solved matrices the participant believed could achieve in Part 2 of Experiment 1, "High Status" is a dummy variable that captures whether the subject was assigned to the high status treatment "Low Ability" is a dummy variable that captures whether the subject was classified as having low ability on the task. $ln(\delta)$ is the estimated dispersion from the mean. Standard errors presented in parentheses. *** denotes significance at the 0.01 level, ** denotes significance at the 0.05 level, * denotes significance at the 0.1 level.

Table 13: Heterogeneous Treatment Effects in Experiment 1

| | (1) Performance | (2) Performance | (3) Beliefs | (4) Beliefs |
|---|---|---|---|---|
| High Status* Performance Part 1 | -3.058*** | -3.058*** | -0.849 | -0.979 |
| | (0.951) | (0.861) | (1.185) | (0.655) |
| High Status | 28.616*** | 28.743*** | 8.794 | 10.195* |
| | (8.868) | (8.102) | (11.001) | (5.983) |
| Performance Part 1 | 2.184*** | 2.392*** | 1.300 | 1.373*** |
| | (0.639) | (0.593) | (0.987) | (0.476) |
| Session Size | | -0.628* | | 0.087 |
| | | (0.358) | | (0.308) |
| Belief practice round | | -0.041 | | 1.513*** |
| | | (0.199) | | (0.137) |
| Medal Fair | | -2.243 | | -0.776 |
| | | (1.555) | | (1.157) |
| Locus of Control | | 0.736* | | 0.381 |
| | | (0.414) | | (0.361) |
| Female | | 3.476** | | -0.426 |
| | | (1.590) | | (1.242) |
| Performed Task Before | | 5.777*** | | 4.793*** |
| | | (1.963) | | (1.682) |
| $ln(\delta)$ | 0.612*** | 0.394 | 0.577*** | -1.000 |
| | (0.260) | (0.287) | (0.206) | (0.636) |
| $N$ | 136 | 136 | 136 | 136 |
| Log-likelihood | -472.418 | -463.537 | -492.051 | -443.101 |

Note: This table presents marginal effects of negative binomial regressions of the model $Performance_i = \beta_0 + \beta_1 PerformancePart1 + \beta_2 HighStatus + \beta_3 PerformancePart1 * HighStatus + Controls'\Gamma + \epsilon_i$, with $\epsilon \sim poisson(\lambda)$. "Performance" is the number of correctly solved matrices in the second part of Experiment 1, "High Status" is a dummy variable that captures whether the subject was assigned to the high status treatment "Performance Part 1' is the number of correctly solved matrices in the first part of the experiment. $ln(\delta)$ is the estimated dispersion from the mean. Standard errors presented in parentheses. *** denotes significance at the 0.01 level, ** denotes significance at the 0.05 level, * denotes significance at the 0.1 level.

relationship between the subjects' performance in the first part of the experiment and the subjects' performance in the second part of the experiment. The estimate associated to "Performance Part I" in Table 13 suggests that for subjects assigned to low status, performance in the parts of the experiment relate positively. For these subjects solving correctly one additional matrix in Part 1 increases their performance in Part 2 by 2.18 tables on average. In contrast, the estimate associated to "Part I* High status" shows that subjects assigned to High Status exhibit a negative relationship between these two parts of the experiment. For these subjects, solving correctly an additional matrix in Part 1 yields a decrease in performance of 3.058 tables on average. The evident asymmetry of these treatment effects accounts for the treatment effects presented in Table 12.

Furthermore, the estimates in Table 13 show that when the subject is given Low Status, solving correctly one additional matrix in Part 1 yields an increase of beliefs in Part 2 of 1.37 average matrices. Moreover, the estimate associated to High Status shows that for any performance level in Part 1, subjects exhibit higher beliefs when assigned to the high status treatments as compared to subjects in Low Status who exhibited the same performance in Part 1.

### D.4 Beliefs by round

To gain further understanding about the influence of social status on thew subjects' beliefs, I investigate the evolution of beliefs over experimental rounds. The aim of this analysis is twofold. First, it seeks to study whether the treatment assignment affects beliefs at the onset of the experiment or whether such influence requires time and experience on the task. Second, it allows me to understand how individuals update beliefs in each round and how the treatment assignment influenced this updating process. The analysis presented in this subsection focuses on low ability subjects, because high ability subjects did not exhibit aggregate belief differences nor performance differences.

Table 14 presents the average beliefs of subjects by round, by treatment, and by ability. I find that low ability subjects assigned to different treatments display similar average beliefs in the first ($U = 0.811, p = 0.417$) and second round ($U = 0.416, p = 0.677$). However, as of the third round, steep treatment differences emerge.[28] Hence, the treatments' influence on the subjects' beliefs emerges after the first rounds. Additionally, the estimates show that the average beliefs by round of low and high ability subjects are statistically indistinguishable when both groups are assigned High Status.[29]

[28]The t-tests of these differences are round 3 ($t(52.046) = -1.819, p = 0.03$), round 4 ($t(59.678) = -1.239, p = 0.110$) and round 5 ($t(61.499) = -1.621, p = 0.055$)

[29]The t-tests of these differences are round 1 ($t(43.284) = 1.206, p = 0.234$, round 2 ($t(60.890) = -0.136, p = 0.891$) ,round 3 ($t(54.178) = 1.001, p = 0.321$), round 4 ($t(60.466) = 0.189, p = 0.850$) and round

Table 14: Performance beliefs by round and by treatment for low ability subjects

| Ability<br>Treatment | Low Ability<br>High Status | Low Ability<br>Low Status | High Ability<br>High Status | High Ability<br>Low Status |
|---|---|---|---|---|
| $\text{Belief}_{r=1}$ | 7.228 | 8.103 | 7.804 | 8.428 |
| | (2.880) | (4.369) | (4.539) | (4.590) |
| $\text{Belief}_{r=2}$ | 8.542 | 8 | 8.464 | 9.463 |
| | (2.582) | (2.449) | (1.971) | (2.079) |
| $\text{Belief}_{r=3}$ | 6.285 | 5.310 | 6.560 | 6.785 |
| | (1.824) | (2.361) | (2.549) | (2.079) |
| $\text{Belief}_{r=4}$ | 4.371 | 3.724 | 4.464 | 4.634 |
| | (2.073) | (2.085) | (1.815) | (2.130) |
| $\text{Belief}_{r=5}$ | 3.3714 | 2.482 | 3.142 | 3.341 |
| | (2.498) | (1.882) | (1.603) | (2.220) |
| $\text{Belief}_{r}$ | 5.96 | 5.524 | 6.257 | 6.345 |
| | (3.029) | (3.549) | (3.285) | (3.259) |

Note: This table presents the averages and standard deviations of beliefs in the second part of Experiment 1 by experimental treatment and round for those subjects classified as low ability. Standard deviations are presented in parentheses.

To understand how subjects update beliefs, I perform a regression relating the subjects' beliefs in a round $r$ to their beliefs and performance in previous rounds. This analysis allows me to distinguish between subjects setting high beliefs to match a high performance level achieved in previous rounds from subjects setting high beliefs due to the influence of the treatment. Evidence supporting the former conjecture would suggest that the high status treatment induced high performance, and that beliefs matched these high performance. Evidence supporting the latter conjecture would corroborate the existence of the proposed mechanism: social status induces internal constraints through beliefs, which consequently affect effort and performance.

The specific statistical model regresses individual beliefs in round $r$ on performance in round $r-1$, beliefs in the previous two rounds, $r-1$ and $r-2$, treatment dummies, and relevant controls.[30] I estimate the model using the Blundell and Bond technique, which has the advantage of allowing the error term of the regression to be correlated with non-observable characteristics of the subjects. Additionally, I instrumented the subject's belief in the previous round, as is typically done in dynamic panel data models, as well as the subject's performance in the previous round.

5 $(t(58.501) = 58.501, 0.661$

[30]This model is the one that best fits the data, additional auto-regressive terms display no statistical significance at the 10% level.

Table 15 presents the estimates of the Blundell and Bond regression. The data suggest that the belief-updating process of low ability subjects differs across treatments in two ways. First, there is a treatment difference in how subjects update beliefs in reaction to achieved performance in the previous round. Specifically, subjects who belong to High Status always updated their beliefs upward with respect to achieved performance in the previous round, while subjects belonging to Low Status do not exhibit this feature. This difference in belief updating suggests that the assignment to High Status induced confidence in the subjects, who believed that their previous achieved could be improved. This result disregards the conjecture of subjects setting high beliefs only to match high performance levels in previous rounds. Instead, subjects in High Status displayed high average beliefs regardless of their performance level achieved in previous rounds.

Table 15: Belief dynamics for low ability subjects

| | (1) | (2) | (3) |
|---|---|---|---|
| Sample | Low Ability | Low Ability/High Status | Low Ability/ Low Status |
| | Belief$_r$ | Belief$_r$ | Belief$_r$ |
| Belief$_{r-1}$ | 0.4170*** | 0.400 *** | 0.342*** |
| | (0.127) | (0.155) | (0.436) |
| Belief$_{r-2}$ | 0.083 | 0.258*** | 0.016 |
| | (0.055) | (0.088) | (0.075) |
| Performance$_{r-1}$ | 0.488*** | 0.521*** | 0.311 |
| | (0.124) | (0.149) | (0.235) |
| Constant | 11.528*** | -1.441*** | -12.201 |
| | (155.444) | (46.090) | (15.728) |
| Controls | Yes | Yes | Yes |
| Observations | 192 | 105 | 87 |
| # instruments | 23 | 23 | 23 |

Note: This table presents estimates of the Blundell and Bond regression of the model $Belief_{ri} = \beta_0 + \beta_1 Belief_{(r-1)i} + \beta_2 Performance_{(r-1),i} + Controls'\Gamma + \epsilon_{r,i}$. "Belief" is a subject's beliefs about the number of correctly solved matrices in round $r$. "Performance" is the number of correctly solved matrices in a round in the second part of the experiment. *** denotes significance at the 0.01 level, ** denotes significance at the 0.05 level, * denotes significance at the 0.1 level.

Second, low ability subjects assigned High Status exhibited stronger state dependence (Heckman, 1981). This means that they exhibit a stronger inertia toward updating beliefs upward as rounds elapse. Subjects in High Status update beliefs upward with respect to their own beliefs in the previous two rounds, while subjects in Low Status also update their beliefs upward, but only with respect to their own beliefs in the previous round. This difference in belief updating between similarly skilled subjects suggests that subjects with high status had the confidence to update beliefs upward and more steeply than their counterparts with low status, and this tendency emerged independently of the subjects' performance in previous rounds.

These two differences in belief updating between low ability subjects support the findings of Eil and Rao (2011) and Mobius et al. (2014), who find that favorable and unfavorable signals about ability are internalized differently. In my experiment, subjects with low ability reacted to an assignment to high status by updating their beliefs upward, displaying optimism about the level of performance they expected to attain in the next round. Such optimism emerged despite them having received unfavorable feedback in the first part of the experiment. In contrast, the low status assignment was ignored by high ability subjects who behaved and formed beliefs as if they were ignoring such unfavorable signals.

### D.5 Regressions post-experimental survey

Table 16: Treatment Effects in Experiment 1

| | (1) Fair Medal | (2) Deserved Medal | (3) Performed Medal | (4) Luck Medal |
|---|---|---|---|---|
| Low Ability | -0.254*** | -0.207* | -0.238*** | 0.050 |
| | (0.119) | (0.109) | (0.094) | (0.088) |
| High Status | -0.241** | -0.200 | -0.270*** | -0.050 |
| | (0.120) | (1.784) | (0.089) | (0.071) |
| Low Ability*High Status | 0.347*** | 0.250* | 0.280*** | 0.020 |
| | (0.174) | (0.154) | (0.119) | (0.117) |
| Constant | 0.634*** | 0.414*** | 0.341*** | 0.121*** |
| | (0.076) | (0.078) | (0.075) | (0.051) |
| N | 133 | 133 | 133 | 133 |
| $R^2$ | 0.090 | 0.038 | 0.090 | 0.01 |

Note: This table presents marginal effects of negative binomial regressions of the model $y_i = \beta_0 + \beta_1 LowAbility + \beta_2 HighStatus + \beta_3 LowAbility * HighStatus + \epsilon_i$, with $\epsilon \sim poisson(\lambda)$. $y_i$ is "Fair Medal" in column (1) a binary variable taking a value of one if the subject believed that the medal assignment was fair, "Deserved Medal" in column (2) a binary variable taking a value of one if the subject believed that the medal assignment was deserved, "Performed Medal" in column (3) a binary variable taking a value of one if the subject believed that the medal assignment was due to performance, and "Luck Medal" in column (4) a binary variable taking a value of one if the subject believed that the medal assignment was due to luck. "High Status" is a dummy variable that captures whether the subject was assigned to the high status treatment "Low Ability" is a dummy variable that captures whether the subject was classified as having low ability on the task. Standard errors presented in parentheses. *** denotes significance at the 0.01 level, ** denotes significance at the 0.05 level, * denotes significance at the 0.1 level.

# Appendix E: Experiment 2

In the second experiment, High Status is exclusively assigned to high ability subjects. The goal of this experiment is to investigate whether this meritocratic assignment maintains the ability differences between high and low ability subjects in the second part of the experiment. Such a result would suggest that the treatment effects in Experiment 1 are entirely caused by the assignment to the treatments and not by other confounding factors. As in the analysis presented in the main body of the paper, I begin showing that the classification of subjects into high and low ability entails considerable performance differences in the first part of the experiment. According to the data, the difference in average performance between low and high ability subjects is of 2.31 standard deviations (Hedge's g, $p < 0.001$).

In the second part of the experiment, high ability subjects outperformed low ability subjects ($U = 2.169, p = 0.015$). This difference is equal to 0.407 standard deviations (Hedge's g, $p = 0.015$ with 1000 bootstrap replications). This result demonstrates that when assigned low status, low ability subjects achieve lower performance levels in the second part of the experiment.

To control for factors other than treatment assignment that could be driving these results, I perform negative binomial regressions of performance in the second part of the experiment on subjects' ability, treatment dummies and relevant controls. The estimates, presented in Columns 1 and 2 of Table 17, confirm the finding that subjects assigned to High Status exhibit higher performance on the task.

Abandoning the binary classification of ability and adopting a continuous measure of ability, i.e. performance in Part 1, leads to findings that are consistent with the aforementioned results. Columns (1) and (2) in Table 18 present the resulting estimates of regressions of Performance in Part 2 on Performance in Part 1 and control variables. I find that higher scores in Part 1 of the experiment lead to higher scores in Part 2. This result shows that the treatment assignment in Experiment 1 fundamentally changed the relationship between subjects' performance in Part 1 of the experiment and their beliefs in Part 2.

Next, I show that unless assigned to High Status, subjects with low ability exhibit lower beliefs as compared to high ability subjects. Table 19 presents the aggregated beliefs of subjects participating in Experiment 2 as well as their beliefs by round and by treatment. On average, the beliefs of high ability subjects are 9.2% higher than those of low ability subjects ($U = 2.99, p = 0.001$).

This result is corroborated by negative binomial regressions of the aggregate measure of beliefs on treatment dummies and relevant controls and regressions of the aggregate measure of beliefs on performance in Part 1 of the experiment and relevant controls. The estimates

Table 17: Treatment Effects in Experiment 2

| | (1) Performance | (2) Performance | (3) Beliefs | (4) Beliefs |
|---|---|---|---|---|
| High Status/High Ability | 2.321* | 2.412* | 3.765** | 4.058*** |
| | (1.362) | (1.299) | (1.719) | (1.549) |
| Group Size | | -0.284 | | -0.556* |
| | | (0.278) | | (0.333) |
| Belief practice round | | 0.289 | | 0.730*** |
| | | (0.198) | | (0.272) |
| Fair Medal | | 1.791 | | 0.937 |
| | | (1.220) | | (1.490) |
| Locus of Control | | 0.035 | | 0.204 |
| | | (0.468) | | (0.530) |
| Female | | -2.003 | | -2.641* |
| | | (1.239) | | (1.596) |
| Task performed before | | 2.479 | | 2.104 |
| | | (1.654) | | (1.455) |
| $ln(\delta)$ | 0.401 | .048 | 0.658 | 0.206 |
| | (0.288) | (0.227) | (0.348) | (0.296) |
| Log Likelihood | -478.157 | -454.599 | -509.192 | -478.792 |
| Observations | 138 | 136 | 138 | 136 |

Note: This table presents marginal effects of negative binomial regressions of the model $y_i = \beta_0 + \beta_1 HighStatus + Controls'\Gamma + \epsilon_i$, with $\epsilon \sim poisson(\lambda)$. $y_i$ is "Performance" in column (1) and column (2) and "Beliefs" in column (3) and column (4). "Performance" is the number of correctly solved matrices in the second part of Experiment 2, "Beliefs' is the number of correctly solved matrices the participant believed could achieve in the second part of Experiment 2, "High Status" is a dummy variable that captures whether the subject was assigned to the high status treatment. $ln(\delta)$ is the estimated dispersion from the mean. *** denotes significance at the 0.01 level, ** denotes significance at the 0.05 level, * denotes significance at the 0.1 level.

Table 18: Heterogeneous Treatment Effects in Experiment 2

| | (1) Performance | (2) Performance | (3) Beliefs | (4) Beliefs |
|---|---|---|---|---|
| Performance Part 1 | 0.908** | 0.817** | 1.633*** | 1.593*** |
| | (0.408) | (0.413) | (0.500) | (0.491) |
| Session Size | | -0.277 | | -0.520 |
| | | (0.293) | | (0.343) |
| Belief practice round | | 0.315 | | 0.778*** |
| | | (0.203) | | (0.249) |
| Medal Fair | | 1.870 | | 1.036 |
| | | (1.280) | | (1.501) |
| Locus of Control | | 0.047 | | 0.212 |
| | | (0.485) | | (0.577) |
| Female | | -2.033 | | -2.732* |
| | | (1.237) | | (1.480) |
| Performed Task Before | | 2.659 | | 2.311 |
| | | (1.743) | | (2.075) |
| $ln(\delta)$ | 1.177*** | 0.0400 | 1.452*** | -0.152 |
| | (0.225) | (0.239) | (0.299) | (0.226) |
| $N$ | 136 | 136 | 136 | 136 |
| Log-likelihood | -459.482 | -454.488 | -485.765 | -477.276 |

Note: This table presents marginal effects of negative binomial regressions of the model $y_i = \beta_0 + \beta_1 PerformancePart1 + Controls'\Gamma + \epsilon_i$, with $\epsilon \sim poisson(\lambda)$. $y_i$ is "Performance" in column (1) and column (2) and "Beliefs" in column (3) and column (4). "Performance" is the number of correctly solved matrices in the second part of Experiment 2, "Beliefs' is the number of correctly solved matrices the participant believed could achieve in the second part of Experiment 2, "Performance Part 1" is the number of correctly solved matrices in the first part of the experiment. $ln(\delta)$ is the estimated dispersion from the mean. Standard errors presented in parentheses. *** denotes significance at the 0.01 level, ** denotes significance at the 0.05 level, * denotes significance at the 0.1 level.

presented in columns (3) and (4) of Table 17 and Table 18 corroborate the result that low ability subjects in the experiment exhibit lower average beliefs and that this is not an artifact of the coarse binary classification of ability. Furthermore, Table 19 show that this difference emerges in the first rounds of the experiment. In particular, high status subjects display higher belief in round two ($U = 3.433, p = 0.001$) and round three ($U = 3.254, p = 0.0003$).

Table 19: Performance beliefs in Experiment 2

| Ability | Low ability | High ability |
|---|---|---|
| Belief$_{r=1}$ | 7.112 | 8.045 |
| | (0.445) | (0.533) |
| Belief$_{r=2}$ | 8.323 | 10.121 |
| | (0.268) | (0.336) |
| Belief$_{r=3}$ | 6.323 | 7.560 |
| | (0.236) | (0.277) |
| Belief$_{r=4}$ | 4.056 | 4.484 |
| | (0.245) | (0.263) |
| Belief$_{r=5}$ | 3.281 | 3.575 |
| | (0.269) | (0.309) |
| Aggregated Beliefs | 29.098 | 33.787 |
| | (1.009) | (1.197) |

Note: This table presents the averages and standard deviations of beliefs in the second part of Experiment 2 by experimental treatment and round. Standard deviations are presented in parentheses.

Subjects in Experiment 2 accessed the rank 1.847 times on average. As in Experiment 1, the frequency of rank-access increased over rounds. After the first round subjects accessed the rank 0.27 times on average and this number surged to 0.94 after the third round of the experiment. This increase in rank-access frequency is driven by more subjects accessing the rank. After the first round 20.3% of subjects who had access to the rank choose to look at it and this number increased to 54,23 % at the end of the experiment.

Across treatments I find no empirical evidence of a difference in rank access behavior with the used measures of rank-access. Specifically, rank access behavior was similar across treatments after the first round ($U = 0.101, p = 0.919$), after the third round ($U = 0.101, p = 0.919$), and over the five rounds experiment ($U = 0.227, p = 0.820$). A regression of the different measures of rank-access frequency on treatment dummies and relevant controls confirm this finding. The regression estimates are presented in Table 20. All in all, these results demonstrate that the lower frequency at which low ability subjects accessed the rank in Experiment 1 was only due to their assignment to the High Status.

Finally, I estimate the statistical models presented in Table 17 with the difference that the different measures of rank access are included in the model and are interacted with the dummy indicated assignment to the High Status treatment. Table 21, presents the regression estimates. I find that high ability subjects, who are also those belonging to high status, exhibit higher performance than low ability subjects when both groups do not access the rank. Also, the estimates of the third and fifth columns show that accessing the rank more often can be detrimental to the performance of high ability subjects. This result suggest that accessing the rank could be costly for these individuals inasmuch as it decreased their available time to perform the task. Nevertheless, such a decrease in performance does not appear in the data for low ability individuals. However, these low skill subjects exhibit lower beliefs as they acquire more information about their ability relative to that of others. This result suggest that the relative performance rank informed these subjects about their ability and they reacted properly to this information by updating their beliefs downwards.

Table 20: Determinants of rank-access in Part 2 of Experiment 2

| | (1) Times Rank Round 1 | (2) Times Rank Round 1 | (3) Times Rank Round 1-3 | (4) Times Rank Round 1-3 | (5) Times Rank All rounds | (6) Times Rank All rounds |
|---|---|---|---|---|---|---|
| High Ability | -0.054 | -0.037 | -0.000 | -0.001 | 0.013 | -0.133 |
| | (0.156) | (0.157) | (0.380) | (0.413) | (0.579) | (0.641) |
| Group Size | | 0.063 | | 0.679 | | 0.837 |
| | | (0.199) | | (0.614) | | (0.922) |
| Fair Medal | | 0.015 | | 0.131 | | 0.440 |
| | | (0.145) | | (0.429) | | (0.710) |
| Locus of control | | 0.024 | | 0.000 | | 0.258 |
| | | (0.060) | | (0.137) | | (0.261) |
| Female | | -0.344 | | -1.124* | | -1.000 |
| | | (0.242) | | (0.598) | | (0.788) |
| Belief practice round | | -0.020 | | 0.023 | | 0.278 |
| | | (0.038) | | (0.132) | | (0.271) |
| Task performed before | | -4.742*** | | -0.913 | | -0.938 |
| | | (1.691) | | (0.669) | | (1.277) |
| $ln(\delta)$ | 0.221 | -0.104 | 1.116*** | 0.995*** | 1.405*** | 1.246*** |
| | (0.564) | (0.730) | (0.375) | (1.140) | (0.286) | (0.344) |
| Observations | 59 | 58 | 59 | 58 | 59 | 58 |
| Log Likelihood | -38.476 | -33.547 | -73.937 | -68.092 | -103.912 | -98.252 |

Note: This table presents marginal effects of negative binomial regressions of the model $TimesRank_i = \beta_0 + \beta_1 LowAbility + \beta_2 HighStatus + \beta_3 LowAbility * HighStatus + Controls'\Gamma + \epsilon_i$, with $\epsilon \sim poisson(\lambda)$. "Times Rank" is the occurrence of rank-access by a subject in the second part of the experiment,"High Status" is a dummy variable that captures whether the subject was assigned to the high status treatment "Low Ability" is a dummy variable that captures whether the subject was classified as having low ability on the task. $ln(\delta)$ is the estimated dispersion from the mean. Standard errors presented in parentheses. *** denotes significance at the 0.01 level, ** denotes significance at the 0.05 level, * denotes significance at the 0.1 level.

Table 21: Treatment effects and rank-access in Experiment 2

| | (1) Performance | (2) Beliefs | (3) Performance | (4) Beliefs | (5) Performance | (6) Beliefs |
|---|---|---|---|---|---|---|
| High Status | 2.648** | 4.284*** | 3.349** | 4.620*** | 3.303** | 4.618*** |
| | (1.326) | (1.595) | (1.343) | (1.632) | (1.343) | (1.668) |
| High Status * Times Rank | -2.419 | -3.349 | -3.639** | -3.009 | -1.624** | -1.528 |
| | (4.792) | (5.028) | (1.449) | (2.097) | (0.816) | (0.989) |
| Times Rank | 0.513 | -1.520 | 0.042 | -1.102*** | 0.058 | -0.649** |
| | (1.475) | (1.212) | (0.364) | (0.381) | (0.221) | (0.267) |
| Group Size | -0.268 | -0.399 | -0.059 | -0.136 | -0.094 | -0.110 |
| | (0.286) | (0.351) | (0.298) | (0.372) | (0.300) | (0.380) |
| Fair Medal | 1.710 | 0.943 | 2.022* | 1.233 | 1.953 | 1.216 |
| | (1.235) | (1.498) | (1.219) | (1.466) | (1.232) | (1.468) |
| Locus of Control | 0.038 | 0.206 | 0.042 | 0.234 | 0.043 | 0.307 |
| | (0.470) | (0.522) | (0.461) | (0.484) | (0.467) | (0.485) |
| Female | -2.084* | -3.149* | -2.501** | -3.812** | -2.356* | -3.606** |
| | (1.214) | (1.624) | (1.241) | (1.637) | (1.243) | (1.599) |
| Belief practice round | 0.299 | 0.740*** | 0.328 | 0.774*** | 0.327 | 0.797*** |
| | (0.199) | (0.276) | (0.201) | (0.277) | (0.201) | (0.277) |
| Task Performed before | 2.400 | 1.661 | 2.074 | 1.343 | 2.019 | 1.301 |
| | (1.669) | (1.528) | (1.730) | (1.551) | (1.713) | (1.553) |
| $ln(\delta)$ | 0.041 | 0.181 | -0.026 | 0.088 | -0.002 | 0.090 |
| | (0.224) | (0.292) | (0.224) | (0.305) | (0.225) | (0.304) |
| Observations | 136 | 136 | 136 | 136 | 136 | 136 |
| Log Likelihood | -454.369 | -477.813 | -452.008 | -474.199 | -452.760 | -474.308 |

Note: This table presents marginal effects of negative binomial regressions of the model $TimesRank_i = \beta_0 + \beta_1 LowAbility + \beta_2 HighStatus + \beta_3 LowAbility * HighStatus + Controls'\Gamma + \epsilon_i$, with $\epsilon \sim poisson(\lambda)$. "Times Rank" is the occurrence of rank-access by a subject in the second part of Experiment 2,"High Status" is a dummy variable that captures whether the subject was assigned to the high status treatment "Low Ability" is a dummy variable that captures whether the subject was classified as having low ability on the task. $ln(\delta)$ is the estimated dispersion from the mean. Standard errors presented in parentheses. *** denotes significance at the 0.01 level, ** denotes significance at the 0.05 level, * denotes significance at the 0.1 level.

# Appendix F: Experimental Instructions

This is an experiment in the economics of decision-making. The instructions are simple and if you follow them carefully and make good decisions, you might earn a considerable amount of money, which will be paid to you via bank transfer at the end of the experiment. The amount of payment that you receive depends entirely on your decisions and effort.

Once the experiment has started, no one is allowed to talk to anybody other than the experimenter. Anyone who violates this rule will lose his or her right to participate in this experiment. If you have further questions when reading these instructions please do not hesitate to raise your hand and formulate the question to the experimenter.

**Part 1**

In the first part of the experiment we will ask you to solve a set of 12 tasks, in each of the tasks you are asked to complete a pattern, to do so, you need to choose among some of the options that we provide. Remember that only one of the options is correct. In this part of the experiment you have 4 minutes in order to complete the set of 12 tasks. With the completion of this task we will place you in one of two groups. At the beginning of this part of the experiment we will ask you to provide a personal goal or target, this is we would like you to estimate how many patterns you would be able to solve in that round. Please provide this goal at your best ability! We would really like to know how accurate your estimates are.

Here is an example, which option do you think is the most accurate to complete the pattern?
(Display Example 1)

Here is another example, which option do you think is the most accurate to complete the pattern?
(Display Example 2)

(Completion set I, programmed to be 5 minutes)

**Treatment assignment**

The following participants have a position in the GOLD group. [Call out ID numbers]. Please come up as we call your name and receive your medal. You will wear your medal for the rest of the exercise. Please remain standing at the front of the room until all medals are distributed.

Let's give the Gold group a round of applause!

**Part 2**

In the second and last part of the experiment you are asked to solve patterns just like the ones that you completed in the first part of this experiment. You need to solve as many patterns as you can, since for each correctly solved pattern you would receive a certain amount of points, which can be exchanged for money at the end of the experiment. Hence the money that you earn in the exercise depends on your performance in this part of the experiment.

(Display only if relative performance ranking available ) While working on the patterns you can also check you ranking with respect to your peers by pressing on the button "check my ranking" located at the bottom of your screen. Mind that this descending ranking only reflects your performance in the task with respect to your peers. You can go back to complete patters by pressing the bottom "Take me back to work".

During this part of the experiment you have 5 rounds, each of 4 minutes, to complete as many patterns as you can. Feedback about your own performance, this is whether you solved correctly a pattern or not, would be given to you as soon as you solved that pattern. A summary of the number of correctly solved and incorrectly solved patterns in the round would be given to you as soon as the round ends. Your final score, this is the amount of points derived from each round, would only be shown to you at the end of the experiment. The exchange rate at which the points can be exchanged for money would be determine is of 0.50 Euro cents per point.

**Ranking access (only for some subjects)** While working on the patterns you can also check you ranking with respect to your peers by pressing on the button "check my ranking" located at the bottom of your screen. Mind that this ranking reflects your performance in the task with respect to your peers in this session. You can go back to work by pressing the button "Take me back to work".

Finally, at the beginning of each round we will ask you to provide a personal goal or target, this is we would like you to estimate how many patterns you would be able to solve in that round. Please provide this goal at your best ability! we would really like to know how accurate are your estimates.

(Completion set II, programmed to take 25 minutes)

**Survey**

(i) With which gender do you identify yourself with?

- Female
- Male
- Other

(ii) How old are you?

(iii) What is your field of study?

(iv) What type of degree are you pursuing at the moment?

- Bachelors
- Graduate
- Other

(v) Have you performed the complete-the-pattern task that you repeatedly solved during this experiment before?

- Yes
- No

(vi) Was the assignment to the gold group deserved?

- Yes
- No

(vii) Was the assignment to the gold group fair?

- Yes
- No

(viii) I usually get what I want in life.

- True
- False

(ix) I need to be kept informed about news events.

- True
- False

(x) I never know where I stand with other people.

- True
- False

(xi) I think that I could easily win a lottery.

- True
- False

(xii) If I do not succeed on a task, I tend to give up.

- True
- False