Meritocracy and Inequality*

Paul-Henri Moisson†

June 20, 2024

Abstract

How do individuals behave in a society that rewards "merit", despite not being all on the same starting line? Does inequality in head starts make meritocracy undesirable? Attempting to answer these questions, this paper develops a model of career concerns in which agents publicly choose among several activities in which to exert effort, and differ along a privately observable characteristic ("head start") that affects their performance. The agents' audience values talent, effort and head start. We highlight two contrasting effects: a displacement effect by which the "poor" (head start-wise) try to avoid a lower talent image and thus avoid the activity chosen by the "rich", and a distinction effect by which the rich try to reap a higher head-start image and thus avoid the activity chosen by the poor. While displacement drags the poor towards activities with lower incentives on effort, distinction pulls the rich towards activities with higher incentives. Interpreting the model in terms of "meritocracy", we emphasize how the dominance of displacement or distinction can cause well-meaning policy interventions to backfire.

JEL numbers: D2, D6, H2, J24, M5.

Keywords: Meritocracy, inequality, image concerns, displacement, distinction.

---

*The author is indebted to Jean Tirole for invaluable guidance and support. The author thanks Robert Gibbons, Johannes Hörner and Charles Angelucci for their numerous and inspiring comments and encouragements. The author thanks Daron Acemoglu, David Austen-Smith, David Autor, Roland Bénabou, Alessandro Bonatti, Olivier Compte, Esther Duflo, Christian Hellwig, Bengt Holmström, Yves Le Yaouanq, Elisa Macchi, Mohamed Saleh, Nikhil Vellodi, Thierry Verdier and audiences at MIT, PSE, TSE, 1st Parisian Behavioral Economics Symposium, SIOE 2023, EEA-ESEM 2023, for helpful discussions and feedback. The author gratefully acknowledges funding from the European Union (ERC, HWS, grant n°101098319)

†Toulouse School of Economics; from September 2024, Paris School of Economics and Ecole des Ponts ParisTech, paulhenri.moisson@tse-fr.eu

1 Introduction

Meritocracy vows to reward "merit", definitions of which generally include a weighted sum of innate talent, acquired abilities and past efforts. The meritocratic vocabulary has been increasingly popular since Young's (1958) famous dystopia.[1] However, and possibly driving this rise, not only does the implementation of meritocracy still generate major policy debates,[2] but the idea of meritocracy itself faces growing and multifaceted criticisms. On the one hand, it is argued that meritocracy, while theoretically desirable, cannot be achieved in practice as any attempt is doomed to be rigged by inequalities in individuals' "head starts", which encompass not only financial wealth, but also human and social capital more broadly. Accordingly, critics have complained that while pretending to reward "merit", meritocracy is in fact rewarding such head starts, and that these either have a low social value – e.g., private lessons aimed at securing a high score on an examination, but not improving the student's long-term productivity –, or are morally undeserved – e.g., when stemming from an individual's social background or genetic characteristics.[3] On the other hand, going back to Young's (1958) original stance, it is argued that "perfect meritocracy" is in fact not desirable because of the inequality it induces, the double punishment it inflicts on the "losers" (adding social stigma to lower material rewards), and, as emphasized more recently, the excessive competition it generates among the "winners" (see Markovits 2019).

Could a unified model reconcile these seemingly contradicting claims? We ask three guiding questions: How do individuals behave in a society that rewards "merit", despite not all individuals being on the same starting line? Does head-start inequality make meritocracy undesirable? What are the policy implications?

We thus study an environment in which agents care about their "merit", as perceived by others (peers, future employers or universities, society, etc.). We mainly follow a positive approach, investigating the consequences of different definitions of "merit". The core of our model is the canonical career-concerns framework (Holmström 1982/1999), which we augment with three key features. Firstly, we allow agents to publicly choose among several

[1]Depending on the definition of "merit", the term "meritocracy" may indeed apply to a wide array of political systems – e.g., from "merit" as academic ability to party loyalty or ideological enthusiasm. For a brief philosophical overview of the notions of *merit* and *meritocracy*, see, e.g., Mulligan (2023).

[2]As two recent illustrations, consider the Education Law of China (2021), and in particular its provisions for private classes and tutoring, or the recent decision by the Supreme Court of the United States in the case *Students For Fair Admissions, Inc. v. President and Fellows of Harvard College* (2023). In both cases, the implementation (and consequences) of "meritocracy" in an unequal environment were among the key issues.

[3]See, e.g., Sandel's (2020) critique of "credentialism". A counterargument claims that, even ignoring head starts and the higher performances they help achieve, the head start-rich deserve their higher status as they tend to exert more effort than the head start-poor – e.g., by pursuing longer and more demanding degrees, or putting up more working hours. Underlining the psychological strength of this argument, recent experimental evidence seems to suggest that when rewarding effort and allocating "merit", individuals do not fully take into account (if at all) the non-merit-based incentives that the agents had to exert effort – e.g., monetary incentives directly linked to performance. See notably Andre (2022).

activities – e.g., academic majors, colleges, jobs, or tasks within an organization –, which in the simplest specification of the model differ only in their precision (inverse of noise variance). Secondly, we introduce a privately observable heterogeneity among agents that affects their (public) performance. We refer to this heterogeneous trait as "head start", whether it stems from financial resources, human capital, social capital, etc. Thirdly, we allow (past) effort and head starts to be valued by the agents' audience – e.g., if effort has a longlasting impact on an agent's productivity, or if head starts embody valuable skills, cultural capital, a taste for studying, etc. Hence, we allow "merit" to be a weighted combination of an agent's (expected) talent, effort and head start.

We show that head-start inequality generates separating equilibria and that, under a standard equilibrium refinement, their structure can be explicitly and uniquely characterized. Driving the separation are the contrasting incentives induced by the audience's weights on talent, effort and head start. When talent image concerns dominate head-start image ones, a *displacement* effect arises: the "poor" (head start-wise) avoid the activity chosen by the "rich" to avoid a lower talent image. Because the negative externality from the rich on the poor's talent image increases with the activity's precision, the poor avoid the rich by moving to activities with lower precision – more generally, lower incentives for effort. Hence, the rich choose an activity with incentives on effort that match the payoff value of expected effort, and displace the poor towards less precise activities. Because the poor thus face lower incentives, they exert less effort. Importantly, the poor would pick the same activity and exert the same effort as the rich if the latter had no (privately observable) head start.

Yet, head-start inequality has another facet. When head-start image concerns dominate talent image ones, a *distinction* effect arises: while the poor would prefer to pool with the rich, it is now the rich who separate from the poor by choosing an activity with even higher precision – more generally, higher incentives on effort –, thereby reaping a higher head-start image, while foregoing the higher talent image they would obtain by pooling with the poor. As a consequence, higher head-start inequality reduces aggregate effort if the displacement effect dominates, and increases it if the distinction effect does. In both cases, head-start inequality drives the rich and the poor apart.[4]

Lastly, the agents' choice of activity depends on their effort image concerns: talent-image concerns generate an incentive to exert effort that increases with activity precision. As a con-

[4]Our vocabulary of "displacement" and "distinction" may require a word of explanation. In a sense, in both cases, one party (the poor or the rich) is *displaced* in that it would choose another activity if the other party were not around, and in both cases too, one party tries to *distinguish* itself from the other party – either the poor signalling their poverty to ultimately signal their talent, or the rich signalling their head start to signal its associated merit value. Our labels stem from the vast sociological literature on "distinction" (e.g., Bourdieu 1979 and following), which focuses almost exclusively on the rich's effort to distinguish themselves from the poor. Once the label "distinction" is attributed to this effect, "displacement" becomes the most natural (remaining) label for the opposed effect.

sequence, absent head-start inequality, agents choose the activity in which the talent-image incentives to exert effort match the weight on effort image – the higher the latter, the higher the agents' favored precision –, thereby tying talent image concerns and effort image ones. Hence, with head-start inequality, displacement arises when effort image concerns dominate head-start image ones, and distinction arises if the opposite holds.

In sum, our three additions to the standard career-concerns model combine as follows: head-start inequality makes agents willing to separate, multiple activities allow them to do so, while the respective weights on talent, effort and head-start images determine how they separate.

To provide an illustration in the context of education, when the displacement effect dominates, students from disadvantaged backgrounds fret about confronting well-prepared or well-connected or highly motivated students in the same educational tracks or institutions, and thus opt for less precise and less rewarding tracks. Conversely, when the distinction effect dominates, well-prepared or well-connected or highly motivated students engage in highly selective and demanding tracks to discourage disadvantaged students from following them, and thus enjoy the reputation attached to their preparation/connections/motivation. Displacement prevails when image weights emphasize effort over head start – e.g., when head starts stem from private lessons aimed at preparing for an examination and thus have little value beyond the examination –, whereas distinction prevails when head starts are praised – e.g., when specific soft skills, cultures or social connections are highly valued by recruiters.

To return to the question of whether head-start inequality makes meritocracy undesirable, we study how the definition of *merit* itself, i.e. the relative weights on talent, effort and head start, affects efficiency. Surprisingly, meritocracy achieves efficiency only if the weights on effort and head start are equal. For, any discrepancy between the two induces distortions in terms of activity and/or effort choices. Such equivalence goes against a responsibility-based view of merit, which would place a zero weight on head starts, and against the interpretation of the model in which head starts are but a privately observed component of talent, which would call for the same weight on talent and head start.

More tentatively, we ask how meritocracy, defined as merit-based image concerns, performs compared with alternative modes of organization. We focus on "spot markets for performance" as a benchmark, in which agents sell their actual performance (or "credentials") to competing buyers. Normalizing the market price of performance to deliver the same reward for effort as meritocracy and focusing on separating equilibria, meritocracy reduces inequality and aggregate effort if displacement prevails, and increases inequality and aggregate effort if distinction does.

Policy. We henceforth take the definition of merit as exogenously given, focusing on the generic case in which the weights on effort and head start differ. Separation, provoked by either displacement or distinction, distorts the agents' activity and effort choices away from efficiency.[5] In addition, the agents' effort may have (positive or negative) externalities on third-parties. These distortions thus raise the question of policy interventions. We distinguish two broad categories of interventions: interventions that affect the "landscape" and interventions that affect the "horizon". The former change the set of activities available to the agents, leaving unaffected the audience's inferences and the agents' subsequent image payoffs, whereas the latter leave the set of activities available to the agents unchanged, but change the agents' image payoffs. We focus on separating equilibria for our policy analysis.

Optimal activity landscape. For a given head-start inequality, what would be the optimal activity characteristics? We assume that the principal presents the agents with a menu of activities to choose from,[6] which differ in their precision and transfers (wages/fees). Agents can alternatively choose an "outside activity", that is beyond the principal's control – e.g., in the context of (national) education, drop out of the schooling system, or go abroad to attend a foreign university. The principal faces the usual trade-off between incentivizing effort and reducing rents. The principal has two means to do so: distorting the activity precision of the party most tempted by the outside option (as standard), or relying on distinction or displacement to relax the participation and incentive constraints. As the magnitude of distinction and displacement increases with head-start inequality, for low head-start inequality, the optimal activity landscape is determined by the comparison between the outside option's precision and the first-best precision, whereas for large head-start inequality it is determined by the comparison between the outside option precision and the weight on head-start image. On a political economy note, varying the welfare weights on the rich and the poor indicates that an oligarchy (interested in the rich's welfare) sets a higher precision for the poor, while a "quasi-Rawlsian" principal (interested in the poor's welfare) sets a lower precision for the rich.

Common policy recommendations. We turn to more limited interventions and highlight their unintended consequences stemming from the displacement and distinction effects. One such popular intervention on the activity landscape is capping activity precision.[7] When

[5] Distortions also arise in pooling equilibria (if any). Indeed, in any such equilibrium, the audience's beliefs depend both on an agent's activity choice and on their performance. As a consequence, in a pooling equilibrium (if any), the agents try to signal their having or lacking a head start through both their activity choice and their performance, which distorts their effort choices.

[6] We interpret these activities as a subset of preexisting activities. Hence, we assume the principal is able to ban all the activities it wants from a (sufficiently large) preexisting set, leaving the agents to choose among the remaining ones, and that the principal is able to set activity-specific transfers.

[7] In France, "selection" in public universities has been opposed by many political leaders and intellectuals over the last few decades. In the United States, proposals to ban some tests for college admission or Sandel's (2020) proposal of a lottery among qualified students for admission to elite colleges can also be interpreted

distinction dominates under *laissez-faire*, a cap on activity precision prevents separation as it "corners" the rich in a lower-precision activity, closer to the poor's *laissez-faire* activity, thereby helping the poor to join them. By contrast, when displacement dominates under *laissez-faire*, separating equilibria survive the introduction of the cap: while the cap dislodges the rich and drives them towards an activity with a lower precision, the poor, who under displacement try to avoid the rich, now migrate towards an activity with even lower precision to avoid the rich. As a consequence, when displacement dominates under *laissez-faire*, both the rich and the poor are strictly worse off with the cap. The unintended consequences of a precision cap under displacement – or conversely, of a precision floor under distinction – suggest that image concerns and head-start inequality can create a "whack-a-mole" game for the policy maker, in which one party (rich or poor) chases the other while circumventing the policy intervention.

Moving away from landscape-changing interventions to horizon-changing interventions, we then investigate (future) income taxation. In our risk-neutral setting, head-start equality implies that the optimal income tax is nil.[8] By contrast, with head-start inequality, the optimal income tax crucially depends on whether displacement or distinction dominates. Indeed, with displacement (resp. distinction), head-start inequality induces a suboptimally low effort by the poor (resp. suboptimally high effort by the rich), which the principal counters with a subsidy (resp. tax). The higher the head-start inequality, the higher the pre-tax distortions and thus the larger the magnitude of the principal's optimal interventions – hence, with our zero benchmark, the higher the subsidy with displacement and the higher the tax with distinction.[9] Furthermore, the same formal analysis delivers insights regarding the optimal intensity of image concerns (or equivalently, the optimal visibility of merit). The higher the head-start inequality, the lower the optimal intensity of image concerns if distinction prevails, but the higher the optimal intensity if displacement does.

Lastly, we consider several important extensions and complements. Allowing for nonlinearities in image payoffs makes the prevalence of displacement or distinction depend not only on the relative payoffs from talent, effort and head start images (as in the linear case), but also on the level of head-start inequality. As an illustration, with (sufficiently) increasing returns to scale for head-start image, displacement prevails for low head-start inequality,

as attempts at curbing precision.

[8] Adding risk-aversion would make it strictly positive. Our insights would then apply starting from this strictly positive benchmark level, rather than from zero.

[9] That the optimal income tax decreases with head-start inequality when displacement prevails can be interpreted as another illustration of the "whack-a-mole" policy game induced by image concerns and head-start inequality. This result can also be related to Rothschild and Scheuer's (2016) analysis of optimal taxation with rent-seeking. In our environment, rents stem from privately observable head starts, which affects the agents' performance and image, and rent-seeking unfolds across activities.

while distinction does for higher head-start inequality.

We study (candidate) pooling equilibria and characterize the different effort incentives they create with respect to separating equilibria. Providing an additional motivation for our focus on separating equilibria, we identify parameter regions for which, with our (running) equilibrium refinement, the only equilibria are separating equilibria.

Introducing a second signalling period – e.g., college after high school, or grad school after undergrad – to capture (some) dynamics of our model reveals that there exist separating equilibria in which the rich and the poor separate not only in their first-period activity choices, thereby revealing their head start, but also in the second period, with the rich choosing again more precise activities. Indeed, because the rich and the poor chose activities with different precisions in the first period, the audience's belief at the start of the second period on the rich's talent is more precise than its belief on the poor's. Hence, for a given second-period precision, a rich agent faces lower incentives to exert effort than a poor agent. Consequently, to ensure that the audience expects them to exert the optimal effort level in the second period, the rich (again) choose an activity with higher precision than the poor.

Another major extension regards the agents' preferences. We show that our main insights are robust to *relative* image concerns, according to which agents compare their payoffs to those of their reference groups (in the spirit of Merton 1957). Relative image concerns deliver interesting additional insights. In particular, they predict that the more a society is segregated along activity lines – i.e., the more individuals compare their payoffs only to those of their activity peers –, then the larger the magnitude of displacement, and the lower the magnitude of distinction.

Outline. Section 2 introduces the model, focusing first on the case of (ex ante) homogeneous agents publicly choosing among several activities, before introducing (privately observable) head starts and investigating the consequences of head-start inequality. It unveils the basic mechanisms and key drivers of the following analyses. It concludes by studying the optimal definition(s) of merit. Section 3 studies (ex post) policy interventions. Section 4 considers several key extensions and complements: nonlinear transfers and image concerns in Section 4.1, activity pooling in Section 4.2, dynamics in Section 4.3 and reference groups and relative image concerns in Section 4.4. Section 5 reviews the literature. Section 6 concludes by briefly evoking several alleys for future research. All proofs are in the Appendix.

2 Model

2.1 No head-start inequality (Homogeneous agents)

There is a continuum of agents, with mass 1. Each agent is characterized by their (unobservable) talent $\theta \in \mathbb{R}$. There is a continuum of activities indexed by $h \in \mathbb{R}_+$, and each agent participates in exactly one activity. Agents may, for instance, be students choosing a major or a college, or they could be prospective workers choosing among job offers from different firms or industries.

After having chosen an activity, each agent chooses an effort level in that activity. The agent's outcome in activity h is then given by

$$y = \theta + e + \varepsilon_h,$$

where $e \geq 0$ is the agent's effort in activity h and ε_h is a random noise, normally distributed with mean zero and variance $1/h$. Hence, activities are indexed by their *precision* $h \in \mathbb{R}_+$. A higher precision corresponds to outcomes more closely related to an agent's talent and effort, e.g., more accurate exams or more efficient monitoring, whereas a lower precision corresponds to a higher role for luck, e.g., due to noisier evaluations or garbled outcomes. As standard in career-concerns environments, activity precision drives the agents' incentives to exert effort. As will be clear shortly, our results still hold in more general environments, e.g., adding outcome-based monetary bonuses, by indexing activities according to their incentives to exert effort. For parsimony and simplicity, we focus on precision.[10]

When choosing activity h and exerting effort e, the agent incurs a cost $g(e)$, where the function g is twice differentiable, positive, strictly increasing and strictly convex, with $g(0) = g'(0) = 0$ and $\lim_{e \to +\infty} g'(e) = +\infty$.

Neither an agent nor the audience (more on the latter shortly) observe the agent's talent θ. All share the same prior, which is normally distributed with mean 0 and precision h_0. Effort is privately observable, whereas activity choices (h) and outcomes (y) are publicly observable by all.

Career/Image concerns. An agent's *audience* may be thought of as embodying (nonexclusively) the other agents, third-parties such as relatives or friends, the agent's supervisors or managers, potential future employers, etc. Each agent values the audience's opinion of them. Namely, the agent cares about a weighted sum of the audience's expectation of their

[10] Moreover, we focus on a sufficiently wide set of activities ($h \in \mathbb{R}_+$) through most of the exposition to rule out corner solutions. We study the latter in Section 3.2.

talent, $\hat{\theta}$, and their effort, $\hat{e}$, given their choice of activity h and outcome y:

$$\psi(h,y) \equiv \hat{\theta}(h,y) + \eta\hat{e}(h,y)$$

with $\eta \in (0,1)$ the weight on the effort image, capturing for instance in the traditional career-concerns view, the long-lasting impact of effort on productivity, or from a moral-desert view, how meritorious effort is.[11]

Each agent thus chooses their activity h and their effort level e to solve:

$$\max_{h\geq 0}\ \max_{e\geq 0}\ \ \mu\psi(h,y) - g(e),$$

where $\mu > 0$ denotes the intensity of image concerns.

We look for Perfect Bayesian equilibria.

Preliminary analysis. The audience updates its belief given the agent's activity choice h and performance y:[12]

$$\hat{\theta}(h,y) + \eta\hat{e}(h,y) = \frac{h}{h_0 + h}(y - e^*(h)) + \eta e^*(h),$$

where $e^*(h)$ is an agent's optimal effort in activity h, and is given by:

$$\frac{\mu h}{h_0 + h} = g'(e^*(h)). \tag{1}$$

In particular, in a given activity, all agents have the same optimal effort level.
Let us define for any h,

$$U(h) \equiv \mu\eta e^*(h) - g(e^*(h)).$$

Let h^* be such that $U(h^*) = \max_{h\geq 0} U(h)$, and thus $h^*/(h_0 + h^*) = \eta$.

Hence, with the above notation, each agent chooses their activity by solving:

$$\max_{h\geq 0}\ U(h).$$

Lemma 1 (Homogeneous agents). *In equilibrium, all agents choose the same activity h^**

[11]For simplicity, we restrict our attention to $\eta < 1$ to avoid corner solutions. Our main insights still obtain if $\eta \geq 1$ (with a finite support for activity precision), or $\eta \leq 0$ (e.g., if effort today damages one's future productivity (say, due to harmful activities), and are robust to heterogeneous η across activities.

[12]The "no-signalling-what-you-don't-know" property implied by PBE (Fudenberg-Tirole 1991) yields that the choice of activity h does not signal anything about θ.

such that $U(h^*) = \max_h U(h)$, *and exert effort* $e^*(h^*)$. *A higher weight on effort* η *induces agents to choose an activity with higher precision.*

2.2 Head-start inequality (Heterogeneous agents)

Suppose now that some agents have a head start but others do not, and that while agents privately know whether they enjoy a head start, the audience does not observe it. We refer to "head starts" in the largest possible meaning, encompassing not only financial means – e.g., ability to pay for private tutoring or comfortable studying conditions at home –, but also human and/or social capital – e.g., soft skills, social connections, taste for studying/exerting effort, intrinsic motivation to perform, etc. Head starts may also be interpreted as a second dimension of "talent", already privately revealed to/learnt by the agents, and independent of the θ-dimension.[13] Our insights hold as long as these head starts are at least imperfectly observable by the audience – a realistic assumption.[14]

For simplicity, suppose that each agent has a head start $w \in \{0, M\}$.[15] Hence, M is a measure of head-start inequality. Headstarts are i.i.d. across agents and independent of talent. Let $p \equiv \mathbb{E}[w]/M \in (0, 1)$ denote the share of the rich in the population. An agent's outcome in activity h when having head start w and exerting effort e now writes as

$$y = \theta + e + w + \varepsilon_h,$$

The audience values a weighted sum of an agent's (individual) talent, effort and head start. Letting $\hat{w}$ denote the audience's expectation of an agent's head start, an agent's weighted image is given by:

$$\psi(h, y) \equiv \hat{\theta}(h, y) + \eta\hat{e}(h, y) + \chi\hat{w}(h, y),$$

with $\chi \in [0, 1]$ the weight on head-start image, capturing for instance how long-lasting and productive head starts are, and/or how meritorious they are.[16]

We look for Perfect Bayesian equilibria. For tractability and clarity, we restrict our attention to equilibria in which the audience has degenerate off-path beliefs, i.e. puts probability

[13] Lastly, head starts can also be interpreted as capturing different (marginal) costs of effort. Our additivity assumption preserves the linearity of the model, thus ensuring tractability.

[14] If the audience receives a signal about the agents' head starts, then our analysis applies conditional on each set of signals suggesting the same distribution of head starts. Our insights continue to hold qualitatively across such signal sets.

[15] Our insights remain unchanged with more than two head-start levels (any finite number or even a continuum). See Appendix C.3 for details.

[16] We restrict our attention to $\chi \le 1$ to ensure the existence of a separating equilibrium for all levels of head-start inequality M (see remark below).

0 or 1 on a deviating agent being rich.[17] In addition, we require that if there exist two agent types $w, w' \in \{0, M\}$ such that, for any (degenerate) off-path beliefs, the differential payoff from deviating to an off-path activity h is higher for a type-w agent than for a type-w' agent (strictly so for some beliefs), then the audience put a strictly higher probability on the deviation coming from a type-t agent than from a type-t' agent.[18] Unless stated otherwise, we restrict our attention to such equilibria and consequently, we henceforth refer to them simply as "equilibria".

We first focus on separating equilibria, referring to Section 4.2 for pooling ones.

Preliminary analysis. In any candidate separating equilibrium, an agent's on-path activity choice reveals their head start, and because the audience thus has degenerate beliefs both on- and off-path, it does not update its beliefs regarding the agent's head start after observing the agent's performance. As a consequence, an agent's optimal effort level in a given activity does not depend on their head start and is still given by (1). Yet, an agent's activity choice now depends on their head start and on the other agents' activity choices.

Namely, in a separating equilibrium, an agent with head start w chooses their activity by solving:

$$\max_{h \geq 0} \left(U(h) + \frac{\mu h}{h_0 + h} w - \mu \left(\frac{h}{h_0 + h} - \chi \right) \mathbb{E}[w|h] \right) \tag{P}$$

Hence, with head-start inequality, an agent's choice of activity has three drivers:

(i) an activity-based, head start-independent incentive, $U(h)$, which absent head-start inequality ($M = 0$) is the sole driver of the agent's choice,

(ii) an incentive stemming from the private benefits of their own head start w, that accrue via their (boosted) talent image, $\frac{\mu h}{h_0+h} w$.

(iii) an incentive stemming from the audience's expectation of the agent's head start and/or the collective impact of their activity peers' expected head starts, $\mathbb{E}[w|h]$: substracting to the agent's talent image their expected head start ($-\frac{\mu h}{h_0+h}\mathbb{E}[w|h]$), while attributing the associated head-start image ($\chi \mathbb{E}[w|h]$).

However, the rich (resp. the poor) also bring a positive (resp. negative) externality on the

[17] Our environment features two signalling stages: activity choice, which is publicly observable, and effort choice, which is privately observable but which influences the publicly observable outcome. These two stages can generate complex interactions between the agents' actions and the audiences' beliefs. In particular, with nondegenerate beliefs, they cause head start-specific distortions in the agents' effort choices within a given activity. See Section 4.2 for details.

[18] This refinement is in the spirit of the D1 criterion, as defined by Banks and Sobel (1987), Cho and Kreps (1987). See Appendix A for details.

poor's (resp. rich's) head-start image, with a magnitude proportional to the weight on head-start image χ.

Therefore, when $h/(h_0 + h) > \chi$, i.e., when talent image concerns dominate head-start ones, the rich are eager to blend with the poor to boost their talent image, but the poor are eager to separate from them to safeguard their own talent image. The opposite holds when $h/(h_0 + h) < \chi$, i.e. when head-start image concerns dominate talent ones: the poor are then eager to blend with the rich to reap the benefits of a high head-start image, while the rich are eager to separate from them to signal to the audience that they are the ones with the strongest motivation/soft skills/social capital/etc. Hence, whenever $h/(h_0 + h) \neq \chi$, one party is eager to separate from the other.

How the rich and the poor separate is determined by the second term in (P), which stems from talent image. Indeed, as an agent's own head start w improves their performance and thus their talent image, it is a complement to activity precision h. Importantly, this complementarity between activity precision and head starts arises endogenously from the agents' image concerns (and signal-jamming attempt). It induces the following sorting condition: The poor separate from the rich by moving towards activities with lower precision, in which the rich's head start is less effective, while on the opposite, the rich separate from the poor by moving towards activities with higher precision, in which their head start is more detrimental to the poor's talent image.

As a consequence, in any separating equilibrium, the poor choose activities with lower precision (more generally, lower incentives on effort) than the rich.[19]

Proposition 1 (**Separating equilibria**). *Absent head-start inequality, the unique equilibrium is all agents choosing activity h^*. By contrast, with head-start inequality ($M > 0$), the unique separating equilibrium is:*

(i) ***(Distinction)*** *If $h^*/(h_0+h^*) < \chi$, the separating equilibrium in which the poor choose activity $h_P = h^*$ while the rich choose activity $h_R > h^*$ where h_R is given by*

$$U(h_R) - \mu\Big(\frac{h_R}{h_0 + h_R} - \chi\Big)M = U(h^*).$$

Hence, h_R strictly increases with M. Moreover, h_R strictly increases with χ and η.

(ii) ***(Displacement)*** *If $h^*/(h_0+h^*) > \chi$, it is the separating equilibrium in which the rich*

[19] As it is clear from Proposition 1, the absence of an upper bound on activity precision $h \in \mathbb{R}_+$ (and $\eta \in (0, 1)$, $\chi \in [0, 1]$) gives the agents enough space to avoid each other, if they want to. We describe in Section 3.2 the consequences of (binding) caps or floors on activity precision.

choose activity $h_R = h^*$ while the poor choose activity $h_P < h^*$ where h_P is given by

$$U(h_P) + \mu\Big(\frac{h_P}{h_0 + h_P} - \chi\Big)M = U(h^*).$$

Hence, h_P strictly decreases with M. Moreover, h_P strictly increases with χ and η.

Strikingly, except in the nongeneric case $h^*/(h_0 + h^*) = \chi$, any head-start inequality $M > 0$ induces the existence of a (unique) separating equilibrium. Specifically, (i) whenever head-start image concerns dominate talent image ones (distinction case), head-start inequality generates separation "upwards" with the rich distancing themselves from the poor to signal their head start, whereas (ii) whenever talent image concerns dominate head-start image ones (displacement case), head-start inequality generates separation "downwards" with the poor avoiding the rich to safeguard their talent image. Put differently, the distinction effect stems from the agents' desire to signal a high head start for the sake of their head-start image itself, whereas the displacement effect stems from the agents' desire to signal a low head start for the sake of their talent image. The higher the head-start inequality M, the further away the agents separate.

Whether distinction or displacement prevails depends on the comparison between the weight on effort image η and the weight on head-start image χ, as $h^*/(h_0 + h^*) = \eta$. Displacement prevails if and only if $\eta > \chi$, while distinction prevails if and only if $\eta < \chi$. In other words, a stronger (resp. milder) emphasis on effort as a component of merit fosters displacement (resp. distinction).

Corollary 1 (**Separating equilibrium payoffs**). *For $\eta \neq \chi$, the unique separating equilibrium,*

(i) The rich's expected payoff is higher than the poor's by an additional term equal to $\mu[h_P/(h_0 + h_P)]M$ with displacement, resp. $\mu[h_R/(h_0 + h_R)]M$ with distinction.

(ii) The difference between the rich's and the poor's expected payoffs increases with head-start inequality M, with the weight on effort image η, with the weight on head-start image χ, and with the intensity of image concerns μ.

While intuitive, the comparative statics in Corollary 1 are worth emphasizing. In the context of education, a higher weight on effort η may stem for instance from a higher long-lasting productivity of effort (e.g., due to higher quality teaching), or from a stronger emphasis on effort as a component of "merit". Similarly, a higher weight on head start χ may stem from more productive soft skills, or more valuable social capital and connections. Then, with head-start inequality ($M > 0$), such increases – either in η or χ – widen the gap between the rich's and the poor's equilibrium payoffs. Intuitively, a higher weight on effort η indirectly disfavors the poor with respect to the rich as they exert less effort than the rich in equilibrium, while a higher weight on head start χ directly disfavors them as they are poor.

2.3 Merit and the desirability of meritocracy

Let us consider some normative implications of our analysis. Before turning to standard policy interventions in Section 3, let us ask two sets of questions. Firstly, how does the definition of *merit* itself, captured by the relative weights on talent, effort and headstart $(1, \eta, \chi)$, affect efficiency and inequality? Secondly and more tentatively, how does *meritocracy*, defined as merit-based image concerns, perform in terms of efficiency and inequality with respect to alternative modes of organization?

Merit, efficiency and inequality. Let us fix the (marginal) social value of effort, denoted by $a \in (0, \mu)$,[20] and measure efficiency by the value of the agents' aggregate effort minus their effort costs, i.e. by the quantity $\mathbb{E}[ae(w) - g(e(w))]$. For generality, we focus on the generic case in which the share of the rich in population p is such that $p \neq 1/2$.[21] What definition of *merit*, i.e. weights (η, χ) keeping the weight on talent normalized to 1, achieves efficiency?

Lemma 2 (Merit and efficiency). *Let $p \neq 1/2$. With head-start inequality ($M > 0$), efficiency is achieved in a pure-strategy equilibrium if and only if $\eta = \chi = a/\mu$.*

Unsurprisingly, efficiency requires the weight on effort to match its social value ($\eta = a/\mu$). More surprisingly, efficiency is achieved only if the weight on effort is *equal to* the weight on head start ($\eta = \chi$). For, any discrepancy between the effort and head-start weights induces either separation in activity (and thus effort) choices, or different effort choices for the rich and the poor despite activity pooling.

Efficiency thus requires effort and head start to stand on an equal footing in the definition of merit. Such equivalence between effort and head start goes against a *responsibility-based* view of merit, according to which only effort would qualify for merit as agents do not choose their talent nor their head start (and thus χ should be equal to 0). Moreover, such equivalence clashes with the interpretation of our model in which head start is a privately-observed component of talent, and should thus be valued as much as the unobserved part of talent (and thus χ should be equal to 1). [We refer to Appendix D.2 for details.]

[20] The social value of effort a may for instance include the direct output from an agent's effort, as well as current and future externalities, long-term productivity improvements (learning-by-doing), etc. We require $a < \mu$ to ensure that effiency can be achieved.

[21] Our result below encompasses both separating and pooling equilibria. While the restriction to $p \neq 1/2$ is irrelevant for separating equilibria, it matters for pooling ones (see Appendix D.1 for details).

Observation 1 (Merit and efficiency). *Efficiency generically requires effort and head start to carry the same weight in the definition of merit. In particular, efficiency requires a strictly higher weight on head start than a responsibility-based view of merit would, but a strictly lower weight than a "privately-observed-talent-component" interpretation of head starts would.*

Notwithstanding, we show in Appendix D.3 that balancing efficiency and inequality reduction leads to a strictly lower weight on head start than on effort, and a weakly lower weight on effort than with strict efficiency: $\chi < \eta \leq a/\mu$.

We henceforth take the definition of merit, i.e. the weights η and χ, as exogenously given – be they determined by the market, employers, universities, society, etc.

Meritocracy vs other modes of organization. How does meritocracy, interpreted as merit-based image concerns, compare with alternative forms of social organization? As a (highly tentative and somewhat naive) comparison, let us consider "spot markets for performance", rewarding agents' actual performance rather than inferred merit. Specifically, let us define "spot markets for performance" as a mode of social organization in which agents sell their actual performance – put differently, sell their *credentials* – to competing buyers.[22] For simplicity, let a_{spot} be the constant marginal value of performance on such spot markets, so that a performance y is worth $a_{spot}y$. [Let us moreover restrict our attention to pure-strategy equilibria and assume that a separating equilibrium is selected whenever it exists.]

Hence, with spot markets for performance, in equilibrium, all agents exert school effort e such that $g'(e) = a_{spot}$, and are indifferent over precision levels. Consequently, they achieve efficiency if and only if a_{spot} is equal to the (marginal) social value of effort. To make things comparable, let us assume that $a_{spot} = \mu\eta$. Then, with respect to such spot markets, meritocracy induces higher aggregate effort if $\eta < \chi$, lower aggregate effort if $\eta > \chi$ and the same level of aggregate effort if $\eta = \chi$.

More interestingly, with such spot markets for performance, in equilibrium the difference between the rich's expected payoff and the poor's is equal to $\mu\eta M$. By contrast, with meritocracy, in the separating equilibria described in Proposition 1, the difference between the rich's and the poor's expected payoffs is equal to $\mu[h_P/(h_0 + h_P)]M < \mu\eta M$ if displacement prevails ($\eta > \chi$), and to $\mu[h_R/(h_0 + h_R)]M > \mu\eta M$ if distinction prevails ($\eta < \chi$). As a consequence, with respect to spot markets for performance, meritocracy heightens inequality if distinction prevails, but mitigates it if displacement does.

[22] Such spot markets for performance or rather *credentials* may echo Sandel's (2020) critique of a contemporary implementation of "meritocracy" that rewards credentials at face value, rather than the *merit* that could be inferred from them. As such, Sandel's critique may be targeted more at a departure away from the meritocratic idea, than at the idea itself.

Lemma 3 (**Meritocracy vs spot markets for performance**). *Compared with spot markets for performance delivering the same marginal reward for effort (i.e. with $a_{spot} = \mu\eta$), meritocracy induces higher aggregate effort and higher inequality if distinction prevails ($\eta < \chi$), but lower aggregate effort and lower inequality if displacement does ($\eta > \chi$).*

In practice however, the existence of such spot markets in which no inference about the agents' types is drawn from their performances (pure credentialism) seems unlikely, and some degree of "meritocracy", in the sense of "merit"-based image concerns, seems inescapable.

3 Policy

Let us investigate policy interventions, taking the definition of merit as given. As shown in Section 2, head-start inequality generates separating equilibria in which the agents' activity and effort choices are distorted away from those that maximize aggregate payoffs. In addition, the agents' effort may have (positive or negative) externalities on third-parties, inducing further distortions away from the socially optimal choices.

Hence, a principal – be it a government, or a (monopsony) firm's executive – may want to intervene. We study several possible interventions. We group them into two main categories: interventions that affect the "landscape" and interventions that affect the "horizon". The former change the set of activities available to the agents, leaving unaffected the audience's inferences and the agents' subsequent image payoffs, whereas the latter leave the set of activities available to the agents unchanged, but change the agents' image payoffs.

We begin by studying in Section 3.1 the optimal activity design (optimal activity landscape) in terms of current transfers and precisions, leaving the structure of future payoffs (wages or images) otherwise unchanged, i.e. equivalently, considering the optimal intervention of a principal able to ban any activity it wants (from a sufficiently large initial set) and design activity-specific transfers, but unable to alter the associated future (image or wage) payoffs. This characterization is of interest both as a theoretical benchmark and for applications in which a principal has such power – e.g., a government on public schools/universities, or an executive on its firm's divisions. We then look at more limited policy interventions. We investigate the impacts of caps (or floors) on activity precision, leaving current transfers and future payoffs otherwise unchanged (Section 3.2), thus considering a principal only able to ban all activities with a precision above or below a certain level. Lastly, switching to interventions that affect the activity horizon, we investigate the optimal taxation of (future) income and the optimal intensity of image concerns (Section 3.3), i.e., considering a principal unable to alter the set of activities currently available to the agents, but able to tax future income or change the visibility of "merit".

We consider a general policy objective, which we refine depending on the application. We define the principal's objective as the weighted sum of (i) the externalities generated by the agents' performance, and (ii) the rich's and the poor's welfare. Namely, let the principal's objective be

$$\max \left(\mathbb{E}_w\big[ae(h(w))^* - \beta_{h(w)}\big] + \big[q_R p W_R + q_P(1-p)W_P\big] \right) \qquad (W_{q_R,q_P})$$

with a the marginal value to the principal of the agents' effort (or equivalently current performance), $q_R, q_P \in [0, q]$ the respective weights on the rich's and the poor's welfare, with $q \in (0, 1)$,[23] and W_R and W_P respectively a rich agent's and a poor agent's expected welfare, and where $h(w)$ denotes the activity choice of an agent with head start w. Hence in particular, the principal's weight on a rich agent's welfare is higher than the one on a poor agent's welfare if $q_R \geq q_P$, and strictly lower otherwise. At the extremes, $q_R = q$, $q_P = 0$ may be interpreted as an *oligarchic* objective, while $q_R = 0$, $q_P = q$ as a *quasi-Rawlsian* objective.

Applications. In the context of education, the principal's objective may have $q_R, q_P > 0$ with either q_R or q_P (or both) equal to q, and $a \geq 0$ (positive externalities from education). By contrast, for an organization's executive interested only in the agents' performance, the objective may have $q_R = q_P = 0$, and a be the (marginal) profit from the organization members' performance.

We focus throughout this Section on separating equilibria.

3.1 Optimal activity landscape

With heterogeneous and privately observable head starts, what does the "second-best" activity landscape look like? We assume that the principal can ban all the activities it wants (from the preexisting set of activities $h \in \mathbb{R}_+$), and design activity-specific transfers (β_h).[24]

Put differently, the principal thus chooses activities' precision $(h_k)_k$ and transfers $(\beta_k)_k$ to maximize the objective W_{q_R,q_P}, subject to the agents' incentive and participation constraints.[25] For simplicity, we assume that the cost of effort $g(\cdot)$ is quadratic: $g : e \mapsto g(e) =$

[23] We take $q < 1$ to take into account a (possibly infinitesimal) cost of public funds, and rule out indifference cases. Namely, denoting by $\lambda > 0$ the principal's marginal cost of public funds, then $q = 1/(1 + \lambda) < 1$, and the marginal value of agents' performance is also normalized by $(1 + \lambda)$ (i.e. $a = \tilde{a}/(1 + \lambda)$).

[24] We focus on the interpretation of "head start" as soft skills and/or human or social capital more generally, and assume that agents face no credit constraint (and thus can pay any fixed fee $\beta_h < 0$, subject to their participation constraint).

[25] To deliver limit results, we allow the principal to offer two activities with the same precisions (with precisions and transfers subject to incentive compatibility, so that the rich and the poor still separate over the two activities).

$e^2/2$, and that $a + \mu\eta \in (0, \mu)$, so that optimal precisions are interior. We focus on the implementation of separating equilibria,[26] assuming that whenever several equilibria coexist, the separating one is selected.

The agents' outside option is another activity, beyond the principal's control, with precision $h_{out} \in \mathbb{R}_+$ and fixed transfer $\beta_{out} \in \mathbb{R}$. In the context of education, the outside option may be dropping out of school, or leaving to study in a foreign university (we study below a case with multiple outside options). For simplicity, we assume that β_{out}, h_{out} are such that it is optimal for the principal to have both the rich and the poor participate.

Formally, the principal thus solves

$$\max_{\left((\beta_{k(w)},\ h_{k(w)})_{w\in\{0,M\}}\right)} \mathbb{E}\big[ae^*(h_{k(w)}) - \beta_{k(w)}\big] + q_R p W_R + q_P(1-p)W_P$$

subject to the participation constraints: for all $w \in \{0, M\}$,

$$\beta_{k(w)} + \mu\eta e^*(h_{k(w)}) - g(e^*(h_{k(w)})) + \mu\chi w \ \geq \ U_{out} + \frac{\mu h_{out}}{h_0 + h_{out}}(w - \mathbb{E}[w|out]) + \mu\chi\mathbb{E}[w|\text{out}],$$

and incentive constraints: for all $w, w' \in \{0, M\}$ such that $w \neq w'$,

$$\begin{aligned}&\beta_{k(w)} + \mu\eta e^*(h_{k(w)}) - g(e^*(h_{k(w)})) \\ &\quad \geq \ \beta_{k(w')} + \mu\eta e^*(h_{k(w')}) - g(e^*(h_{k(w')})) + \mu\Big(\frac{\mu h_{k(w')}}{h_0 + h_{k(w')}} - \chi\Big)(w - w').\end{aligned}$$

We refer to the *first-best precision level* h^{FB} as the one that maximizes the principal's objective absent head-start inequality, subject only to the agents' participation constraint. It is given by $\mu h^{FB}/(h_0 + h^{FB}) = a + \mu\eta$.

The payoff from the outside option depends on an agent's head start and on the audience's beliefs about the head start of agents choosing the outside option. We say that incentives are *aligned* if a deviation to the outside option is attributed to a poor agent, and *countervailing* if it is attributed to a rich agent.

To build the intuition, let us first describe two polar cases.

Imprecise outside option ($h_{out} < h^{FB}$) and aligned incentives. As an illustration, in the context of education, once outside of the educational system, an agent cannot send any signal about their academic ability, and thus $h_{out} \ll h^{FB}$. With aligned incentives, the binding constraints are the poor's participation constraint and the rich's incentive constraint. As a consequence, indexing by R the rich's activity and by P the poor's, the second-best precision

[26] We assume that the principal can offer activities with the same precision, but different transfers, such that the rich and the poor separating over the two options is incentive compatible.

levels h_R^{SB} and h_P^{SB} are given by

$$\begin{cases} \dfrac{\mu h_R^{SB}}{h_0 + h_R^{SB}} = a + \mu\eta, \\ \dfrac{\mu h_P^{SB}}{h_0 + h_P^{SB}} = \max\left(\dfrac{\mu h_{out}}{h_0 + h_{out}},\ a + \mu\eta - (1 - q_R)\dfrac{p}{1-p}M\right) \end{cases} \tag{2}$$

The rich's rent is equal to $\mu[h_P^{SB}/(h_0 + h_P^{SB}) - h_{out}/(h_0 + h_{out})]M$, strictly increases with the poor's precision and may be nonmonotonic with M. The rich exert the first-best effort level, whereas the poor exert an effort below the first-best level as their activity's precision is distorted downwards to reduce the rich's rent. The higher the weight on the rich's welfare, the lower the distortion.[27]

Highly precise outside option ($h_{out} > h^{FB}$) and countervailing incentives. In the context of education again, the principal may be facing competition from highly selective foreign universities. With countervailing incentives, the binding constraints are now the rich's participation constraint and the poor's incentive constraint. Assuming that solutions are interior, the second-best precision levels h_R and h_P are given by

$$\begin{cases} \dfrac{\mu h_R^{SB}}{h_0 + h_R^{SB}} = \min\left(a + \mu\eta + (1 - q_P)\dfrac{1-p}{p}M,\ \dfrac{\mu h_{out}}{h_0 + h_{out}}\right), \\ \dfrac{\mu h_P^{SB}}{h_0 + h_P^{SB}} = a + \mu\eta. \end{cases} \tag{3}$$

The poor's rent is equal to $\mu[h_{out}/(h_0+h_{out}) - h_R^{SB}/(h_0+h_R^{SB})]M$, strictly decreases with the rich's precision and may be nonmonotonic with M. The poor now exert the first-best effort level, whereas the rich exert a strictly higher effort as their activity's precision is distorted upwards to reduce the poor's rent. The higher the weight on the poor's welfare, the lower the distortion.[28]

General case. Whether the principal chooses aligned or countervailing incentives depends not only on the difference between the precision of the outside option h_{out} and the first-best precision h^{FB}, but also on head-start inequality M and on the weight on head-start

[27] The difference $h_P^{SB}/(h_0 + h_P^{SB}) - \chi$, and thus the (magnitude of the) displacement or distinction effects does not appear in the second-best precision levels. Yet, they influence the transfers $\beta_R^{SB}, \beta_P^{SB}$.

$$\begin{cases} \beta_R^{SB} = g(e^*(h_R^{SB})) - \mu\eta e^*(h_R^{SB}) + \mu\left(\dfrac{h_P^{SB}}{h_0 + h_P^{SB}} - \chi\right)M + U_{out}, \\ \beta_P^{SB} = g(e^*(h_P^{SB})) - \mu\eta e^*(h_P^{SB}) + U_{out}. \end{cases}$$

[28] The transfers $\beta_R^{SB}, \beta_P^{SB}$ are now given by

$$\begin{cases} \beta_R^{SB} = g(e^*(h_R^{SB})) - \mu\eta e^*(h_R^{SB}) + U_{out}, \\ \beta_P^{SB} = g(e^*(h_P^{SB})) - \mu\eta e^*(h_P^{SB}) - \mu\left(\dfrac{h_R^{SB}}{h_0 + h_R^{SB}} - \chi\right)M + U_{out}. \end{cases}$$

image χ. Indeed, the principal can rely on the distinction effect to reduce the rich's rent under aligned incentives, and on the displacement effect to reduce the poor's rent under countervailing incentives. Head-start inequality determines the magnitude of these potential gains: for low head-start inequality, displacement/distinction have little traction and the principal resorts to (standard) precision distortions to reduce rents, while by contrast, for large head-start inequality, displacement/distinction have a strong hold on the agents' choices and the principal relies on them to reduce rents.

Proposition 2 (**Optimal activity design, exogenous outside option**). *Suppose the principal can choose the transfers β_k and precision h_k of activities $k \in \{R, P\}$ to implement separating equilibria. Then,*

(i) The higher the welfare weight of the rich (q_R), the higher the poor's precision, and the higher the welfare weight of the poor (q_P), the lower the rich's precision.

(ii) For sufficiently low head-start inequality M, the principal chooses aligned incentives if $h_{out}/(h_0 + h_{out}) < a + \mu\eta$, and countervailing incentives if $h_{out}/(h_0 + h_{out}) > a + \mu\eta$. By contrast, for sufficiently large head-start inequality M, the principal chooses aligned incentives if $\chi > h_{out}/(h_0+h_{out})$, i.e. if distinction prevails in the outside activity, and chooses countervailing incentives if $\chi < h_{out}/(h_0 + h_{out})$, i.e. if displacement prevails in the outside activity.

Succinctly, two take-aways from Proposition 2 are that: (i) with respect to a principal putting equal welfare weights on rich and poor agents ("formal equality"), an "oligarchy" sets a higher precision for the poor, while a "Rawlsian" principal sets a lower precision for the rich; (ii) the principal's choice of aligned or countervailing incentives is determined by the comparison between the outside option's precision and (a) for low head-start inequality, the weight on effort (image weight and externalities $a + \mu\eta$), (b) for large head-start inequality, the weight on head-start image (χ).[29]

Discussion. We conclude this Section with two remarks focusing on the consequences of competition in labor markets.

Endogenous outside options: A distant, competitive labor market. Suppose that the agents' outside option is now a distant, competitive labor market.[30] Accessing that market

[29]The same insights obtain if the agents' images are zero-sum from the principal's perspective, implying milder optimal incentives.

[30]We assume that in a competitive labor market, competing firms offer a menu of incentive-compatible contracts specifying a precision and a transfer, and we focus on the separating equilibria described in Proposition 1. In addition, we assume that firms in that distant labor market do not react to our principal's activity design – e.g., because the principal's target population is sufficiently small with respect to the one of the distant labor market.

entails a fixed transportation cost $d > 0$ to the agents. There, competing firms make a profit $a.e$ from employee effort e, and offer activities with precision $h_{out,R}$ and $h_{out,P}$ where $h_{out,R}, h_{out,P}$ are as described in Proposition 1, replacing $\mu\eta$ by $a+\mu\eta$. With such endogenous outside options,

(i) (Displacement) If $a + \mu\eta > \mu\chi$, then for any level of head-start inequality, h_R^{SB} and h_P^{SB} are given by (2), replacing h_{out} by $h_{out,P}$.

(ii) (Distinction) If $a + \mu\eta < \mu\chi$, then for any level of head-start inequality, h_R^{SB} and h_P^{SB} are given by (3), replacing h_{out} by $h_{out,R}$.

Competitive vs monopsonistic labor markets. Let us assume that the outside options are endogenously determined as we have just described, and let us investigate the consequences of labor market competition – e.g., firms competing for workers, or universities competing for students.[31] Hence, if displacement prevails ($a + \mu\eta > \mu\chi$), a monopsony induces a higher effort from the poor (and the same effort from the rich) than competitive labor markets, whereas if distinction prevails ($a + \mu\eta < \mu\chi$), it induces a lower effort from the rich (and the same effort for the poor) than competitive labor markets – strictly so if either head-start inequality M is low, or displacement prevails and the rich are few (p low), or distinction prevails and the poor are few ($1 - p$ low). Put differently, competitive labor markets induce larger distortions in the agents' (activity and effort) choices than a monopsony.[32] Therefore, in terms of aggregate welfare, a (duly disciplined) monopsony dominates a competitive labor market – strictly so if either head-start inequality is low, or displacement prevails and the rich are few, or distinction prevails and the poor are few.

3.2 Precision caps

We illustrate in a simple setting the (unintended) consequences of imposing a cap on activity precision. Such a cap may stem from lowering precision (adding "noise") in the most precise activities, from outright bans on precise activities – e.g., removing specific fields from school curricula –, or from making an activity irrelevant – e.g., making participation in specific ability tests or extracurricular activities (say, music or sports) irrelevant for university admission/recruitment decisions. [Symmetrically, we consider below the consequences of precision floors – e.g., in the education context, adding a mandatory base examination, or

[31] For a formal comparison, suppose that in any activity h, agents receive a transfer $ae^*(h)$, which may be either a wage or a tuition fee depending on the application, and let the policy weights q_R, q_P be such that $q_R = q_P = 0$, corresponding to standard profit maximization for firms. (See Section 4.1 for additional details.)

[32] These results may be compared with those of Bénabou and Tirole (2016). In their setting with an observable and a non-observable task (focusing as we do on the least-cost-separating outcomes), competition leads to higher distortions than monopsony in the allocation of effort between the two tasks, yet also to higher incentives for effort.

in an industry context, introducing a minimum monitoring/surveillance technology.]

Let us thus assume that imposing a cap $\overline{h}$ on activity precision amounts to banning all activities with precision $h > \overline{h}$.

Proposition 3 (**Precision caps: Equilibrium characterization**). *Suppose that the principal sets a precision cap $\overline{h} < h_R$ with h_R the activity chosen by the rich under laissez-faire. Then, with the precision cap $\overline{h}$,*

(i) ***(Distinction)*** *If $h^*/(h_0 + h^*) < \chi$, there exists no separating equilibrium.*

(ii) ***(Displacement)*** *If $h^*/(h_0 + h^*) > \chi$ and $\overline{h}/(h_0 + \overline{h}) > \chi$, there exists a unique separating equilibrium: the rich choose activity $\overline{h}$ and the poor choose activity $h_P(\overline{h})$ such that*

$$U(h_P(\overline{h})) = U(\overline{h}) - \mu\Big(\frac{h_P(\overline{h})}{h_0 + h_P(\overline{h})} - \chi\Big)M.$$

With distinction, a cap on precision is effective at "cornering" the rich and enabling some poor to join them – the lower the cap the more so. By contrast, with displacement, while the cap forces the rich into an activity with lower precision, they further displace the poor towards an activity with an even lower precision. In the context of education, removing (or adding noise to) an examination can thus result either in the poor catching up and competing in the same tracks as the rich when effort image matters less than head-start image (e.g., when soft skills and social connections matter more than school effort), or the poor being displaced towards even less precise tracks when effort image matters more than head-start image (e.g., when effort has higher long-term productive impacts than head starts).

Analogous insights hold with a precision floor (e.g., introducing a mandatory examination), which destroys all separating equilibria under displacement, but generates a separating equilibrium with higher precisions for all agents under distinction. In a sense, image concerns and head-start inequality trigger a "whack-a-mole" policy game whereby the poor and the rich keep escaping from/chasing the other party, circumventing the principal's policy goal – regardless of whether the principal's "hammer" is an activity ban or, as we will study next, income taxes.

Corollary 2 (**Precision caps: Separating equilibrium payoffs under displacement**). *Suppose the principal sets a precision cap $\overline{h}$ between the rich's and the poor's laissez-faire activity precisions ($h_P < \overline{h} < h_R$). Suppose $h^*/(h_0 + h^*) > \chi$ (displacement). Then, in the (unique) separating equilibrium under the cap, both the poor and the rich are strictly worse off than in the (unique) separating equilibrium absent the cap.*

Hence, when displacement prevails and separation persists (before and after the intervention), a cap on precision makes both the poor and the rich worse off. The impact of a precision cap on the principal's objective W_{q_R,q_P} further depends on how much it values the agents' effort. Setting a precision cap strictly reduces aggregate effort if displacement prevails and separation persists, but may increase it if distinction does.

3.3 (Future) income taxation and intensity of image concerns

We briefly study income taxation, emphasizing a striking property of the optimal income tax in our environment. We leave a detailed study of optimal taxation (and redistribution) in our image-concerns environment for future work.

The principal can commit. We assume that taxation is "activity-blind" and thus that taxes and transfers cannot be conditioned on (past) activity choice or characteristics – e.g., because organizations (be they universities or firms) are able to masquerade their line of business whenever it is in their interest to do so for tax purposes. We restrict our attention to a linear tax on income and assume that the principal has a zero marginal cost of funds.

We focus on the career-concerns interpretation of our model, in which an agent's image is their expected future wage. A (persistent) income tax thus applies to the agents' future income ($\psi(h, y)$). (As throughout Section 3, we restrict our attention to separating equilibria.)

With a linear income tax τ, an agent's optimal effort in activity h, $e^*(h, \tau)$, is given by

$$g'(e^*(h,\tau)) = (1-\tau)\frac{\mu h}{h_0 + h}.$$

An agent with head start w choosing an activity with precision h has an expected payoff before redistribution given by

$$(1-\tau)\mu\Big(\frac{h}{h_0+h}(w - \mathbb{E}[w|h]) + \eta e^*(h,\tau) + \chi\mathbb{E}[w|h]\Big) - g(e^*(h,\tau)),$$

contributing tax proceeds

$$\tau\mu\Big(\frac{h}{h_0+h}(w - \mathbb{E}[w|h])] + \eta e^*(h,\tau) + \chi\mathbb{E}[w|h]\Big)$$

to the principal.

The principal redistributes all tax proceeds to the agents in a lump-sum fashion (see remark below on performance-based redistribution). Hence, with separation, an agent with head start w choosing an activity with precision h has a post-redistribution utility equal to

$$(1-\tau)\mu\Big(\frac{h}{h_0+h}(w-\mathbb{E}[w|h])+\eta e^*(h,\tau)+\chi\mathbb{E}[w|h]\Big)-g(e^*(h,\tau))+\tau\mu(\eta\mathbb{E}[e^*(h,\tau)]+\chi pM)$$

where the last expectation is taken over (all) agents' activity choices.

Let $U(h,\tau)\equiv(1-\tau)\mu\eta e^*(h,\tau)-g(e^*(h,\tau))$, and denote $h^*(\tau)\equiv\arg\max_h U(h,\tau)$. Hence, for all τ, the precision h^* that maximizes $U(\cdot,\tau)$ is such that $h^*/(h_0+h^*)=\eta$. The next characterization follows from Proposition 1.

Lemma 4 (**Equilibrium characterization, income tax**)**.** *Let $\tau<1$ be the linear income tax rate. Then, absent head-start inequality, the unique equilibrium has all agents choosing activity h^*. By contrast, with head-start inequality ($M>0$), the unique separating equilibrium is:*

(i) ***(Distinction)*** *If $\eta<\chi$, the separating equilibrium in which the poor choose activity h^* while the rich choose activity $h_R(\tau)>h^*$ where $h_R(\tau)$ is given by*

$$U(h_R(\tau),\tau)=U(h^*,\tau)+(1-\tau)\mu\Big(\frac{h_R(\tau)}{h_0+h_R(\tau)}-\chi\Big)M.$$

In addition, if g' is (weakly) concave, $h_R(\tau)$ strictly increases with τ.

(ii) ***(Displacement)*** *If $\eta>\chi$, it is the separating equilibrium in which the rich choose activity h^* while the poor choose activity $h_P(\tau)<h^*$ where $h_P(\tau)$ is given by*

$$U(h_P(\tau),\tau)=U(h^*,\tau)-(1-\tau)\mu\Big(\frac{h_P(\tau)}{h_0+h_P(\tau)}-\chi\Big)M.$$

In addition, if g' is (weakly) concave, $h_P(\tau)$ strictly decreases with τ.

As a consequence, absent head-start inequality ($M=0$), the principal implements the first-best effort level with the income tax $\tau^{FB}=0$.[33] By contrast, with head-start inequality $M>0$ and whenever $\eta\neq\chi$, a (unique) separating equilibrium exists, in which by Lemma 4, if the agents' cost of effort is quadratic (g' linear), the higher the income tax, the further apart the rich and the poor separate in terms of precision (the larger $|h_R-h_P|$). Intuitively, a higher income tax "smoothes the landscape" by flattening activity characteristics/incentives, thereby making both parties more mobile across activities. Hence, to mitigate the impact of higher head-start inequality, should the optimal tax decrease with head-start inequality to make both the rich and the poor less mobile and reduce the distortions in activity choices?

[33]The zero optimal income tax stems from the risk-neutrality of agents. Adding risk aversion would yield a strictly positive optimal tax, to which the distortions we evidence below would add. We maintain the risk-neutrality assumption for simplicity.

For simplicity, we assume that the agents' effort has no externalities ($a = 0$), and so the principal's objective W_{q_R,q_P} is a weighted sum of the rich's and the poor's payoffs. For $q_R = q_P = q$, the principal thus solves

$$\max_{\tau} \left[p\Big(\mu\eta e^*(h_R(\tau)) - g(e^*(h_R(\tau)))\Big) + (1-p)\Big(\mu\eta e^*(h_P(\tau)) - g(e^*(h_P(\tau)))\Big)\right]$$

Our main insight follows from the equilibrium characterization of Lemma 4.

Proposition 4 (**Income taxation**). *Suppose the principal places equal weights on the rich's and the poor's welfares ($q_R = q_P$). Suppose the agents' cost of effort is quadratic.*[34] *Then, absent head-start inequality, the optimal income tax is nil, whereas with head-start inequality ($M > 0$), assuming that agents play the separating equilibria described in Lemma 4,*

(i) ***(Distinction)*** *If $\eta < \chi$, the optimal income tax is strictly positive and strictly increases with head-start inequality M.*

(ii) ***(Displacement)*** *If $\eta > \chi$, the optimal income tax is strictly negative (i.e. a subsidy) and strictly decreases with head-start inequality M.*

The sign and monotonicity of the optimal income tax with respect to head-start inequality thus depend on whether distinction or displacement prevails. With displacement (resp. distinction), head-start inequality induces a suboptimally low effort by the poor (resp. suboptimally high effort by the rich), which the principal counters with a subsidy (resp. tax). The higher the head-start inequality, the higher the pre-tax distortions and thus the larger the magnitude of the principal's optimal interventions, i.e., the higher the subsidy with displacement and the higher the tax with distinction.

Remark: Education. In the context of education (e.g., students in high school or college), students' image concerns may stem mostly from being subsequently admitted to a high-quality college or graduate school. One may assume that the funding of (private but also public) colleges or universities increases with the "quality" of its students, and that the larger the funding, the higher the quality of the education they can deliver. Hence, in a (stylized) setting with competing colleges or universities, this Section's analysis may also apply to a tax on these colleges' or universities' funding. The above results then suggest that whether universities should be taxed or subsidized (and to what extent) depends on the magnitude of head-start inequality, and on whether displacement or distinction prevails.

[34] Our results hold more generally for a twice continuously differentiable marginal cost of effort g' if g' is (weakly) concave and $g''(0) = 0$ in the case of distinction ($\eta < \chi$), and if g' is (weakly) convex in the case of displacement ($\eta > \chi$). [These conditions are sufficient, but not necessary in general.]

Remark: Performance-based redistribution. While the principal is activity-blind, it observes current performance (which may also be interpreted as current income) and could thus condition the redistribution of tax proceeds on the agents' performances. Let us consider two polar cases in which the principal redistributes the tax proceeds (collected over the whole population) only among the agents who achieve a performance: (a) above a threshold, (b) below a threshold (in a lump-sum fashion among those agents).[35] Performance-based redistribution affects the agents' incentives: case (a) induces higher effort, whereas case (b) induces lower effort. Hence, when displacement prevails, redistribution among the agents who achieve a performance above a threshold (case (a)) may improve efficiency, yet worsen inequality, whereas when distinction prevails, redistribution among the agents who achieve a performance below a threshold (case (b)) may improve efficiency and reduce inequality.[36]

Remark: Optimal intensity of image concerns (μ). Departing from (future) income taxation, the principal may be able to engineer/influence the intensity μ of the agents' image concerns – e.g., either by affecting their time horizon/discounting factor (in the career-concerns interpretation of the model), or by changing the publicity of the agents' "merit" (in the social status/image concerns interpretation of the model). As an illustration, let us assume that the principal's value from the agents' effort does not depend on μ (and is strictly positive). Then, the same analysis as above yields that the optimal intensity of image concerns depends on whether distinction or displacement prevails. Namely, with head-start inequality, when distinction prevails, optimal image concerns are *less intense* (lower μ) than absent head-start inequality, while when displacement prevails, they are *more intense* (higher μ) – in both cases, the more so the higher the head-start inequality. In terms of publicity of agents' merit, the higher the head-start inequality, the lower the optimal publicity if distinction prevails, but the higher the optimal publicity if displacement does.

[35] Case (a) arises for instance when higher-achieving agents (or agents with a higher current income) have an exclusive access to publicly-funded goods – see e.g., Fernández and Rogerson (1995) for a model in which access to education is only partially publicly provided and thus entails private costs, and agents are credit-constrained. In Fernández and Rogerson (1995), the exclusion of poorer agents from the redistribution of tax proceeds stems from credit constraints (e.g., the lower the tax rate, the lower the public funding of education, and thus the more excluded from redistribution the poor). By contrast, in our environment, it is both their having no head start *and* the separation induced by head-start inequality that makes the poor less likely to benefit from redistribution, as the poor separate in an activity with lower effort incentives than the rich. All else being equal, the larger the head-start inequality, the more excluded the poor.

[36] In our environment, the exact magnitude of the additional incentives coming from performance-based redistribution depends nontrivially on the parameters. Indeed, absent performance-based redistribution, the distribution of the poor's performances has a lower mean than the rich's (due to their having no head start and exerting lower effort), but also a lower precision. As a consequence, whether the marginal effort incentive generated by performance-based redistribution is higher or lower for the poor or the rich – and by how much – depends on the parameters. Deriving the optimal threshold for redistribution is thus a nontrivial problem, which we leave for future work.

4 Extensions

4.1 Transfers and nonlinear image concerns

In the workhorse environment of Section 2.2, for any head-start inequality $M > 0$, the prevalence of distinction or displacement depends only on the difference between η and χ. We introduce in this Section a more general model, showing that nonlinearities in current transfers or in image concerns make the prevalence of distinction or displacement depend on the level of head-start inequality M.

Let us assume that in addition to their image concerns, each agent receives a direct transfer $\beta_h \in \mathbb{R}$ (e.g., current wage or tuition fee), which depends on the activity $h \in \mathbb{R}_+$ chosen by the agent and the audience's (equilibrium) beliefs. As before, we restrict our attention to separating equilibria. Hence, with such transfers, each agent now chooses their activity h and their effort level e to solve:

$$\max_{h\geq 0} \max_{e\geq 0} \Big(\beta_h + \mu\psi(h, y) - g(e)\Big).$$

Assumption 1. *Let the transfer in activity h be a function of the expected effort (with degenerate beliefs), $\hat{e}(h) = e^*(h)$, and expected head start $\hat{w}(h) \in \{0, M\}$:*

$$\beta_h \equiv b(e^*(h)) + c(\hat{w}(h)),$$

with $b : \mathbb{R}_+ \to \mathbb{R}$ a continuously differentiable function, and $c : \mathbb{R}_+^* \to \mathbb{R}_+$. Define for any h,

$$U(h) \equiv b(e^*(h)) + \mu\eta e^*(h) - g(e^*(h))$$

Assume that $U(h)$ is continuously differentiable, first strictly increasing then strictly decreasing, with a unique interior maximum. Let h^* be such that $U(h^*) = \max_{h\geq 0} U(h)$.

Briefly put, Assumption 1 requires that transfers be additively separable in expected effort and head start, and single-peaked with respect to expected effort. In particular, the shape of $U(h)$ satisfies Assumption 1 whenever the transfers $b(e^*(h))$ are a weakly increasing and weakly concave function of $e^*(h)$. The precision h^* (still) strictly increases with the weight on effort η.[37]

We say that transfers exhibit increasing (resp. decreasing) returns to scale in head-start inequality if $c(M)/M$ increases with $M > 0$ (resp. decreases with M).

To make things concrete, let us consider two polar (and naive) applications:

[37] In addition, if the transfers $b(e^*(h))$ are monotonic with the effort $e^*(h)$, then $h^*/(h_0+h^*) \geq \eta$ if $b(e^*(h))$ increases with $e^*(h)$, resp. $h^*/(h_0 + h^*) \leq \eta$ if $b(e^*(h))$ decreases with h.

(i) *Wages bid by firms competing on the labor market:* Firms' profit is a strictly increasing function of their employees' effort and head start. Hence, $b(\hat{e})$ and $c(\hat{w})$ increase with $\hat{e}$ and $\hat{w}$.[38]

(ii) *Tuition fees set by universities competing on the student market:* Suppose that universities receive funding from the government or from alumni as an increasing function of their students' future productivity/wages/social prestige, and thus as an increasing function of their students' expected effort and head start. As larger government funding or larger donations from alumni allow a university to lower its fees, competition in the student market implies that $b(\hat{e})$ and $c(\hat{w})$ increase with $\hat{e}$ and $\hat{w}$ (i.e. that fees decrease with the students' expected effort and head start).[39]

Proposition 5 (**Transfers**). *Suppose Assumption 1 holds. Then, Lemma 1 and Proposition 1 hold, except for the comparative statics of* h_R, h_P*, with* $U(\cdot)$ *and* h^* *as defined in Assumption 1 and replacing* χ *by* $\chi + c(M)/(\mu M)$*. In particular, if transfers exhibit increasing (resp. decreasing) returns to scale in head-start inequality* M*, then there exists* $M^* \in [0, +\infty]$ *such that displacement prevails for* $M < M^*$*, while distinction does for* $M > M^*$ *(resp. distinction prevails for* $M < M^*$ *and displacement does for* $M > M^*$*).*

Our analysis of nonlinear transfers extends to nonlinear image concerns (generalizing Assumption 1), as long as transfers and image concerns remain additively separable in the different components of merit (talent, effort, head start) and linear with respect to talent.[40]

Interestingly, with nonlinear transfers or image concerns, the dominance of either displacement or distinction depends in general on the level of head-start inequality. As an illustration, suppose transfers/images have first increasing then decreasing returns to scale in head-start inequality – e.g., if starting from zero, additional increments to the rich's head start are first increasingly valuable as they bring the rich closer to a critical threshold, but decreasingly valuable past the threshold. Such a "step-like" shape can be expected in par-

[38] We assume that firms (or universities, see below) compete to attract workers (or students) and are able to offer (incentive-compatible) menus of precisions h and wages (or tuition fees) β_h.

[39] In the case of government funding, universities may receive funding that increases with their students' expected effort, but *decreases* with their expected head start – e.g., if the government aims at reducing inequality and vows to handicap head start-rich students. Then, $c(\hat{w})$ may decrease with $\hat{w}$ – e.g. as a higher (average) head start may either *require* higher fees to make up for lower State funding, or *allow* higher fees as students expect higher future wages.

[40] Our insights further hold when transfers are not additively separable in expected effort and expected head start. In particular, when transfers are given by a strictly concave function of expected performance, distinction and displacement still lead to separating equilibria when image concerns are sufficiently intense (μ high) and the weight on head-start image (χ) either sufficiently high or sufficiently low. We refer to Appendix H.1 for details. Intuitively, strict concavity with respect to performance induces the poor to prefer a higher effort level under separation than the rich (as the rich's performance builds on their head start, they face a lower marginal incentive to exert effort). This creates an additional incentive for the poor to pool with the rich. As a consequence, for separating equilibria to survive, image concerns must be sufficiently intense and the weight on head-start image sufficiently high or sufficiently low.

ticular when agents compete for a limited number of prizes, with success in such a contest consisting in achieving a score sufficiently above those of rivals. In such an environment, displacement may prevail for both low levels and high levels of head-start inequality, with distinction prevailing only for intermediate levels.

This observation bears additional implications when there are more than two head-start levels in the population. Suppose for instance that there exists a "middle class" with head start $m \in (0, M)$. With linear image payoffs, either displacement prevails between all head-start levels, or distinction does. By contrast, if transfers and/or image concerns exhibit (sufficiently) increasing returns to scale in expected head start, then, for m sufficiently low and $M - m$ sufficiently high, displacement prevails between the poor and the middle class, while distinction prevails between the rich and the middle class.[41]

Remark: The desirability of meritocracy (continued). We noted in Section 2.3 that the desirability of meritocracy with respect to alternative modes of organization can depend on whether displacement or distinction dominates. Our analysis of nonlinear transfers and image concerns implies that the desirability of meritocracy may thus further depend on the level of head-start inequality.

We henceforth resume our main environment, as outlined in Section 2.2, with no transfers and linear, additively-separable image concerns.

4.2 Pooling equilibria

Let us consider candidate pooling equilibria and more generally, the consequences of the audience having nondegenerate beliefs. With pooling (or nondegenerate beliefs), head starts are imperfectly revealed by activity choices. Hence, the audience updates its beliefs on an agent's head start after observing their performance, which creates an incentive for the agents to signal their (having or lacking a) head start via their performance. In other words, while agents always "signal-jam" their talent via their performance (alone), they signal their head start *only* via their activity choice when the audience has degenerate beliefs, whereas they signal their head start via *both* their activity choice *and* their performance when the audience has nondegenerate beliefs.

[41]To make things formal, suppose for instance that there are no transfers, that the effort image payoff is given by $\mu\eta\hat{e}$ (as in the linear case), but that there exists $m^* > 0$ such that the head-start image payoff is nil for $\hat{w} \in [0, m^*]$, and equal to $\mu\chi(\hat{w} - m^*)$ for $\hat{w} > m^*$. Then, for $\chi > \eta$ and $m < m^* < M - m$, displacement prevails between the poor and the middle class, while distinction prevails between the rich and the middle class.

An agent with head start $w \in \{0, M\}$ now chooses their activity h and effort e by solving

$$\max_h \max_e \mathbb{E}_{\theta,\varepsilon_h}\Big[\mu\eta\hat{e}(h,y) - g(e) + \frac{\mu h}{h_0+h}\Big(e + w - \hat{e}(h,y) - \hat{w}(h,y)\Big) + \mu\chi\hat{w}(h,y)\Big]$$

where $y = \theta + e + w + \varepsilon_h$. As $\hat{e}$ and $\hat{w}$ now depend not only on h but also on y, the optimal effort within a given activity depends in general on an agent's head start and on the audience's prior belief.

Lemma 5 (**Nondegenerate beliefs and optimal efforts**). *Fix an activity $h \in \mathbb{R}_+$. Fix the audience's prior belief $p(h) \in (0,1)$ that an agent choosing activity h is rich. If interior and continuous with respect to $p(h)$, the optimal effort levels in activity h of a rich agent and of a poor agent are both strictly higher than $e^*(h)$ if $h/(h_0+h) < \chi$ (distinction region), and both strictly lower than $e^*(h)$ if $h/(h_0+h) > \chi$ (displacement region).*

As intuitive, when head-start image is the dominant concern (distinction), agents try and signal their having a head start by achieving a higher performance, and to this end, they exert a higher effort. Symmetrically, when talent image is the dominant concern (displacement), agents try and signal their lacking a head start by achieving a lower performance, and to this end, reduce their effort.

Whether it is the rich or the poor who distort their effort the most with respect to $e^*(h)$ depends on the audience's prior belief. In particular, the optimal effort level of a rich agent in activity h lies further away from $e^*(h)$ than the optimal effort level of a poor agent if $p(h)$ is in a neighborhood of 0, and closer to $e^*(h)$ if $p(h)$ is in a neighborhood of 1. The effort distortion is most pronounced for the agents whom the audience expects the least.

We henceforth resume our focus on degenerate off-path beliefs for the audience.

Proposition 6 (**Pooling equilibria with degenerate off-path beliefs**). *Let $p \equiv \mathbb{E}[w]/M$.*
(i) If $\chi = \eta$ and $p \neq 1/2$, there exists a unique pooling equilibrium in pure strategies: All agents choose activity h^* such that $h^*/(h_0+h^*) = \eta$ and exert effort $e^*(h^*) = (g')^{-1}(\mu\eta)$.
(ii) If $\chi \neq \eta$ and p is close to 0 or close to 1, there exists no pooling equilibrium in pure strategies.

Part (ii) in Proposition 6 thus provides an additional motivation for our focus on separating equilibria: for $\chi \neq \eta$ and p in a neighborhood of 0 or in a neighborhood of 1, the only equilibria in pure strategies are the separating equilibria described in Proposition 1.

4.3 Dynamics

Since in our one-period framework, separation in activity choices perfectly reveals the agents' head starts, one may wonder whether separation could persist over time – or appear in the first period to begin with. We thus sketch the dynamics of our environment in a two-period framework. The two stages may for instance describe two stages in one's education – e.g., high school and college, or undergraduate and graduate school. For simplicity, we assume that the agents' head starts remain constant over the two periods – e.g., because periods are short and head starts are nonfinancial.

We show that in this two-period environment, separation can obtain in the first period, thereby perfectly revealing each agent's head start from the beginning. Secondly, after agents separate in the first period and while (on path) no privately observable heterogeneity remains in the second period, separation persists as the rich choosing a higher precision than the poor in the first period makes the audience's end-of-period-1 belief about their talent more precise, which leads the rich to prefer a higher precision in period 2 than the poor.

Let $\delta \in (0, 1)$ denote the discounting factor across the two periods. In period 1, agents first choose an activity, indexed by its precision h_1, then an effort level e_1 in that activity, which entails a period-1 cost of effort $g(e_1)$. Then, their period-1 performance y_1 is realized and publicly observed. In period 2, agents again first choose an activity h_2 and then an effort level e_2, incurring their period-2 cost of effort $g(e_2)$. Their period-2 performance y_2 is realized and publicly observed. Hence, the set of information I about a given agent available to the audience at the end of period-2 is given by $I = \{h_1, y_1, h_2, y_2\}$. An agent's image concern is the weighted sum of their expected talent, sum of expected (past) efforts and expected head start given I – e.g., in a competitive recruitment environment, an agent's image concern is the discounted sum of their future wages.

An agent with publicly observable activity choices and outcomes $I = \{h_1, y_1, h_2, y_2\}$ thus maximizes[42]

$$\mathbb{E}\left[\delta\mu\Big(\mathbb{E}[\theta|I] + \eta(\mathbb{E}[e_1|I] + \mathbb{E}[e_2|I]) + \chi\mathbb{E}[w|I]\Big) - g(e_1) - \delta g(e_2)\right]$$

We refer to *fully separating equilibria* as equilibria in which the rich and the poor separate in both periods. In any fully separating equilibrium, the audience's end-of-period-2 belief on an agent's talent θ with activity choices h_1, h_2 has precision $h_0 + h_1 + h_2$. By linearity, for a given choice of activities h_1, h_2, an agent's second-period effort level $e_2^*(h_1, h_2)$ in any fully

[42] In contrast to Holmström (1982), we assume that the image payoff only occurs at the end of the second period, which we see as a more consistent assumption in the context of education. As a consequence, as we show below, in equilibrium, the second-period effort will be higher than the first-period one, i.e. effort increases over time, whereas in Holmström (1982), the existence of a first-period image payoff implies that effort decreases over time. Adding an image payoff at the end of the first period would leave our insights unchanged.

separating equilibrium is thus given by

$$g'(e_2^*) = \frac{\mu h_2}{h_0 + h_1 + h_2}. \tag{4}$$

For simplicity, we assume that $\eta \in (0, 1/2)$, which will yield interior solutions.[43]

To ensure uniqueness of the equilibrium absent head-start inequality, we assume passive beliefs about effort, i.e., that the audience only uses the agent's period-1 precision choice (h_1) to form its beliefs about e_1.[44]

With these assumptions, for a given period-1 activity choice h_1, an agent's period-2 on-path activity choice $h_2^*(h_1)$ is such that $g'(e_2^*) = \mu\eta$, i.e.

$$\frac{h_2^*(h_1)}{h_0 + h_1 + h_2^*(h_1)} = \eta,$$

and in particular, the period-2 on-path activity choice $h_2^*(h_1)$ is a strictly increasing function of h_1. In addition, for activity choices $h_1, h_2^*(h_1)$, an agent's period-1 effort level $e_1^*(h_1)$ is given by

$$g'(e_1^*) = \frac{\delta\mu h_1}{h_0 + h_1 + h_2^*(h_1)}. \tag{5}$$

Hence, let h_1^* be defined by

$$\frac{h_1^*}{h_0 + h_1^* + h_2^*(h_1^*)} = \eta,$$

and thus in particular, $h_2^*(h_1^*) = h_1^*$, so that $h_1^*/(h_0 + 2h_1^*) = \eta$.

Proposition 7 (**Equilibrium characterization, two-period model**). *Suppose agents choose among a continuum of activities indexed by their precision $h \in \mathbb{R}_+$, and that their cost of effort is quadratic, with $g(e) = e^2/2$. Then, absent head-start inequality, the unique equilibrium in pure strategies is all agents choosing activity h_1^* in period 1 and activity $h_2^*(h_1^*)$ in period 2. By contrast, with head-start inequality ($M > 0$), there exists a fully separating equilibrium in pure strategies:*

(i) ***(Distinction)*** *If $2\eta < \chi$, in period 1 the poor choose activity h_1^* while the rich choose*

[43] The upper bound 1/2 comes from a normalization to keep the total weight on effort, 2η, strictly below 1.

[44] A rationale is that the audience forms its beliefs using only the realized history at the time the agent took their decision. Hence, in particular, the audience does not update its beliefs about e_1 after observing the individual's period-1 performance (y_1), period-2 precision choice (h_2) and period-2 performance (y_2). Put differently, the audience believes that the set of strategies considered by the agent (on- and off-path) is such that for any two strategies (h_1, e_1, h_2, e_2) and (h_1', e_1', h_2', e_2'), if $h_1 = h_1'$, then $e_1 = e_1'$.

activity $h_R > h_1^*$ *where* h_R *is given by*

$$[\delta\mu\eta e_1^*(h_R) - g(e_1^*(h_R))] = [\delta\mu\eta e_1^*(h_1^*) - g(e_1^*(h_1^*))] + \delta\mu\Big(\frac{h_R + h_2^*(h_R)}{h_0 + h_R + h_2^*(h_R)} - \chi\Big)M,$$

while in period 2*, the poor choose activity* $h_2^*(h_1^*)$ *and the rich choose activity* $h_2^*(h_R) > h_2^*(h_1^*)$*.*

(ii) ***(Displacement)*** *If* $2\eta > \chi$*, in period* 1 *the rich choose activity* h_1^* *while the poor choose activity* $h_P < h_1^*$ *where* h_P *is given by*

$$[\delta\mu\eta e_1^*(h_P) - g(e_1^*(h_P))] = [\delta\mu\eta e_1^*(h_1^*) - g(e_1^*(h_1^*))] - \delta\mu\Big(\frac{h_P + h_2^*(h_P)}{h_0 + h_P + h_2^*(h_P)} - \chi\Big)M.$$

while in period 2*, the rich choose activity* $h_2^*(h_1^*)$ *and the poor choose activity* $h_2^*(h_P) < h_2^*(h_1^*)$*.*

In other words, Proposition 7 shows that with initial head-start inequality, there exist equilibria with immediate and persistent separation in activity choices. Separation measured either by the distance between activity precisions, or by the one between effort levels, decreases over time – in fact, in the second period both the rich and the poor exert the same effort $((g')^{-1}(\mu\eta))$.

4.4 Relative image concerns

While in our baseline specification, agents care about the *absolute* levels of their images, we now investigate an alternative specification in which they care about their *relative* levels. Relative image concerns capture the positional property of prestige concerns, while also being consistent with career-concerns interpretations in which an agent's image is their future wage but the utility the agent ultimately derives from their future wage depends on how the latter compares to the others, or in which an agent's chances of being promoted depend on their relative "qualities" with respect to their rivals'.[45]

In the spirit of Merton (1957), we distinguish two *reference groups* for each agent: the whole society and the agent's activity peers.The agents' weights on each reference group may stem from society's division along (or mobility across) activity lines. we emphasize that milder across-activity image concerns – e.g., due to a more divided, less mobile society – tend to foster displacement.

Formally, suppose that an agent's payoff in activity h and with performance y is a weighted sum of their *local*, within-activity image $\psi(h, y) - \mathbb{E}[\psi(h, y)|h]$, and their *global*,

[45]See in particular Frank (1985), and for more recent investigations, Langtry (2022) and Butera et al (2022).

across-activity image $\psi(h,y) - \mathbb{E}[\psi(h,y)]$:

$$(1-\zeta)\Big(\psi(h,y) - \mathbb{E}[\psi(h,y)|h]\Big) + \zeta\Big(\psi(h,y) - \mathbb{E}[\psi(h,y)]\Big) \tag{6}$$

where $\zeta \geq 0$ captures the relative weight of the agent's global image. [Our previous specification thus corresponded to purely across-activity image concerns ($\zeta = 1$), while for $\zeta = 0$, image concerns are purely within-activity: agents only compare themselves and/or are only compared to their activity fellows.]

The extent to which image concerns are across- or within-activity may depend in particular on the extent to which activities, careers or parts of society more generally are clustered. For instance, as a side-product of exerting effort in a given activity, agents may learn activity-specific knowledge – either technical or relative to a profession's/firm's cultural and social norms. If such knowledge is valuable only in that given activity and worthless in others, image concerns may be mostly within-activity as agents expect future competition mostly from their activity fellows. By contrast, if there is no such activity-specific knowledge, image and career concerns are across-activity as agents expect future competition from other agents across activities. Similarly, if society is clustered along activity lines, so that agents in different activities have few interactions – e.g., living and working in different neighbourhoods, having different lifestyles –, images may be mainly within-activity as agents put a higher weight on their comparisons with respect to their activity-peers.

Proposition 8 (Relative image concerns and reference groups). *All formal results in Sections 2.1 and 2.2, namely Lemma 1, Proposition 1 and Corollary 1 hold with relative image concerns (as defined by (6)), replacing η by $\zeta\eta$, and χ by $\zeta\chi$.*

Let us study the comparative statics of the displacement and distinction effects with respect to ζ.[46] Indeed, as a society becomes more divided along activity lines – e.g., as education or work organization become more specialized and students or workers in different activities fields interact less –, across-activity comparisons may matter less (ζ may decrease). By contrast, as a society becomes more mobile – or at least more transparent as agents can more easily observe the lifestyle of other agents –, across-activity comparison may be heightened (ζ may increase).

Corollary 3 (Distinction and displacement in divided vs mobile societies). *The more mobile the society (the higher ζ), the larger the magnitude of distinction and the smaller*

[46] Absent transfers, the optimal precision h^* is such that $h^*/(h_0 + h^*) = \zeta\eta$, and thus the regions of displacement and distinction do not depend on ζ: displacement prevails if $\eta > \chi$, while distinction does if $\eta < \chi$.

the magnitude of displacement.[47] *Conversely, the more divided the society (the lower ζ), the smaller the magnitude of distinction and the larger the magnitude of displacement.*

5 Related literature

Theoretical literatures. This paper builds on several theoretical literatures, too vast to be summarized here. The founding lineage is the literature on career concerns, initiated by Holmström's (1982) seminal contribution from which the core of our model is borrowed. Within this literature, Dewatripont, Jewitt and Tirole (1999a, 1999b) investigate the role of activities' information structure,[48] allowing for complementarities between talent and effort, and applying their analysis to a multitasking environment.[49] Closely related to our main application, MacLeod and Urquiola (2015) consider a model in which students first exert effort to get admitted to a college, then exert effort in college, before going on a competitive job market. In their model, college choice is observable (but test results for college admission are not), and colleges vary in their selectivity, which influences the audience's expectation about the distribution of skills and efforts of the students they admit. These works differ from ours in that they consider *ex ante* identical agents, whereas the key driver of our analysis is an (initial) privately observable heterogeneity, which distorts the agents' equilibrium effort choices, generating displacement or distinction.[50] A notable exception is Madsen, Williams and Srkzypacz (2022), who consider a two-activity environment (a "safe" one and a "risky"

[47] Namely, the higher ζ, the higher the rich's precision h_R under distinction, and the higher the poor's precision h_P under displacement.

[48] Relatedly, while most of the literature on career concerns assumes that agents' performances are observable throughout the employment relation, Bonatti and Hörner (2017) consider an environment in which only breakthroughs are observed, yielding in particular that wages are single-peaked over time (conditional on no breakthrough being observed).

[49] Cisternas (2018) introduces strategic skill acquisition, studying environments in which effort is a direct input both to current production and to skill acquisition, and finds that the audience's uncertainty on whether to attribute a higher output to new skills or to noise can lead to suboptimally low effort. In our setting, the audience's ex ante uncertainty regarding an agent's head start – and the agents' strategic reaction to this uncertainty – leads to suboptimally low effort (from the poor) when displacement prevails, but suboptimally high effort (from the rich) when distinction does.

[50] Our investigation of activity choices with image concerns further relates to the literature on endogenous group formation with peer effects. In particular, our model can be compared with Bénabou (1993). In Bénabou (1993), agents choose their skills and location, while in ours they choose effort and activity. In Bénabou (1993), positive externalities from high-skill neighbors make would-be high-skill workers willing to pay more to live in a high-skill neighborhood, and the limited availability of land then generates segregation. By contrast, in our model, we rule out congestion in activities, but positive (or negative) externalities from peers' head starts make the rich or the poor willing to incur a higher or a lower precision, hence providing a suboptimal effort, and segregation obtains when the chased party can escape sufficiently far away (in terms of precision) from the chasing party. In a different vein, Board (2009) considers peer effects alone, and emphasizes that "private provision" of activities leads to excessive segregation, while Staab (2022) adds status concerns to peer effects and shows that, with private provision, status concerns mitigate the segregation induced by peer effects. In our setting, peer effects (image externalities) can either make the rich willing to blend with the poor and the poor willing to avoid them, or the other way around, while within-activity "status concerns" can arise from the agents' relative image concerns with respect to their activity peers (see Section 4.4). As opposed to Staab (2022), in our setting, the complementarity between the privately observable heterogeneity and activity precision arises endogenously via the signal-jamming attempt of the agent (talent image concerns).

one) and agents with unobservable talent, yet with a privately-observable, activity-specific head start. A head start increases an agent's probability of success in the risky task,[51] and is useless in the safe task. The authors study the optimal incentive scheme, combining monetary bonuses and promotions. Hence, in contrast to their environment, in ours head starts are useful in all activities, which generates the distinction/displacement chase across activities.

The second literature on which this paper builds is the signalling literature, starting with Spence (1973). In our model, agents try to signal their having or lacking a head start, and the complementarity between the agents' head start and their precision choices, which generates the sorting condition, arises endogenously from the agents' effort to influence their talent image ("signal-jamming" induced by talent image concerns). Our agents' trade-off between talent, effort and head-start images can be compared with studies of signalling to multiple audiences with imperfectly aligned preferences, such as Austen-Smith and Fryer (2005). In their setting, agents choose a one-dimensional variable to signal a privately observable two-dimensional type: the alignment between the two dimensions of an agent's type is determined in equilibrium via the opportunity cost of underinvesting in one dimension. By contrast, in our environment, agents know only one dimension of their type (head start) and face a single audience to which they send a two-dimensional signal (activity choice and performance): the alignment between the three dimensions of the agent's image is determined in equilibrium via the opportunity cost of choosing suboptimal effort incentives (precision) – either too high or too low.

Our policy analysis contributes to the (already rich) study of optimal incentives with career/image concerns, in the wake of seminal contributions such as Gibbons and Murphy (1992). Our study of optimal income taxation echoes Rothschild and Scheuer's (2016) study of optimal taxation with rent-seeking, where in our environment, rent-seeking stems from privately observable head starts and unfolds across activity choices. In addition, our results on the optimal intensity of image concerns are related to Ali and Bénabou (2020), who find that the optimal "visibility" of prosocial behavior solves a trade-off between incentivizing effort and revealing societal preferences. Likewise, in our setting, the optimal visibility of "merit" solves a trade-off between incentivizing effort and increasing the distortions in the agents' activity choices, which goes towards higher visibility when displacement prevails, resp. lower visibility when distinction does.

This paper studies the relations between the allocations of "merit" and material rewards in society, and it is thus related to contributions comparing different forms of (social) or-

[51] In addition, head starts and talent are assumed to be complementary in the agent's probability of success, whereas in our model they are substitutes in the agents' performance.

ganization, such as Coase (1937), Green and Stokey (1983), Cole, Mailath and Postlewaite (1992), Burdett and Coles (1997), and Fernández and Galí (1999) who compare contracts and contests, or markets and contests, underlining that contests induce excessive effort with respect to markets, but achieve a higher matching efficiency (strictly so when agents face borrowing constraints). In our model, meritocracy has contrasted consequences with respect to alternative forms of organization. As mentioned in Section 2.2, with respect to "spot markets for performance", it induces lower effort and higher payoff inequality if displacement prevails, but higher effort and lower payoff inequality if distinction does.[52] The matching efficiency of meritocracy – the accuracy of the audience's beliefs about the agents' talent – is higher when distinction prevails than when displacement does.

Divergent behaviors for the rich and the poor, as in our model, have been given many explanations. An important literature, pioneered by Arrow (1973), relies on self-fulfilling beliefs by which agents either imperfectly observe the characteristics of different activities (as in Piketty 1995, Alesina and Angeletos 2005, Bénabou and Tirole 2006), or face different audience expectations regarding their effort or the causes of their success/failure (as in Coate and Loury 1993, Piketty 1998). These environments stand in contrast to ours in which activity parameters are perfectly known, and agents face *ex ante* identical expectations from the audience. Accordingly, the policy implications of our model differ.

Empirical literatures. A vast empirical literature – not limited to economics – describes how students' backgrounds affect their choices, as well as the role of expectations and narratives. Inspiring our work are several seminal contributions from sociology. In particular, Bourdieu and Passeron (1970) and Bourdieu (1979) present and analyze sociological evidence of the separating outcomes we label as "displacement" and "distinction", in particular in the context of education. In addition to objective head starts (as in our model), they identify as an additional driver of separation the narratives, promoted by some elites, discouraging "lower-class" individuals from choosing more selective and demanding education tracks. Boudon (1973), building on Merton's (1957) notion of "reference groups", provides another explanation for the same outcomes, based on class-specific aspirations and beliefs.[53] In our model, the agents' differentiated choices obtain even in the absence of any such narratives

[52]In the wake of Baker, Gibbons and Murphy (1994), investigating the interplay of such spot markets with merit-based image concerns may yield interesting insights.

[53]In a related vein, Müller (2022) provides empirical evidence of the strong impact of "parental pressure" on children's education choices, interpreted as including both coercion and transmission of the parents' beliefs and preferences. Our model may suggest an alternative explanation of these findings: the parents' reaction to their child's prospective application to a given university may reveal to the child how much parental support (material and immaterial) they could expect were they to attend that university. Such parental support constitutes a head start (privately observable and affecting the student's performance). While these two explanations point to the same outcome, they call for different remedies.

or class-specific aspirations.

Closer to our framework, Bourdieu and Saint-Martin (1970) analyze the relative image weights of talent, effort and head start (to rely on our model's vocabulary) in French education, notably emphasizing the joint depreciation of effort and appreciation of talent and head start (across fields and across educational tracks). As an illustration, they provide evidence that if literature and mathematics are perceived as the most prestigious fields by both professors and students, it is partly because they are considered as those in which effort is the least useful and in which talent or head start (and specifically the latter in the case of literature, e.g., having read, or heard of many books not mentioned in school programs) matter more than effort exerted at school.[54]

Within the economics literature, Burzstyn, Egorov and Jensen (2019) provide experimental evidence of students' fear of revealing their ability to their peers, which induces students with lower grades to decline opportunities for additional preparation for the SAT and additional diagnostic tests – a pattern consistent with the displacement effect. In the same experiment, the authors show that, depending on a school's social norm ("smart-to-be-cool" vs "cool-to-be-smart", which we capture in our model with the image weights on talent, effort and head start), higher weights on talent and effort induce agents to sign up for additional preparation and diagnostic tests, i.e. selecting a higher precision – as predicted by our model.[55]

Some implementations of the policy interventions we study have been empirically documented. In particular, Moreira and Pérez (2022) provide a rich analysis of the consequences of the introduction of competitive exams to select certain federal employees (following the 1883 Pendleton Act), which may fit in our model as raising the precision of this career track. Moreira and Pérez (2022) find that the exams left the share of upper-SES applicants unchanged, increased the share of middle-class applicants and decreased the share of lower-SES applicants. From our model perspective, this may suggest that displacement prevailed between middle-class and lower-SES applicants (the higher precision making joining the civil service more attractive to the middle class and less attractive to lower SES, for the sake of talent signal-jamming), and that distinction prevailed between middle-class and higher-SES

[54] Coupled with our model, their analysis thus suggests that distinction should prevail in these fields. However, they also describe literature as a much less *precise* field than mathematics (in terms of performance evaluation), both by intentional design and because of its intrinsic nature (e.g., the ambiguities of definitions and evaluations of "style"), and thus literature may not enable the same magnitude of distinction, if any, as mathematics.

[55] Relatedly, Ashraf, Bandiera and Lee (2014) study visibility interventions in a nationwide health worker training program in Zambia. They find that higher "employer recognition" and "social visibility" increase performance – consistently with a higher μ in our model. In addition, they find that raising "social comparisons", i.e. increasing the weight on within-activity image concerns lowers performances, in particular for "low-ability trainees" – consistently with a displacement effect and our predictions in Section 4.4 regarding the impact of strong within-activity image concerns.

applicants (the higher precision thus making the higher-SES applicants only slightly more inclined to pool with middle-class applicants, as higher-SES applicants remain predominantly concerned with signalling their head start). The coexistence of displacement between lower-SES and middle-class applicants and of distinction between higher-SES and middle-class ones is consistent with head starts having increasing returns to scale (see Section 4.1).[56]

Empirical studies outside of the education context may also be interpreted in the light of our model. As an illustration, Bursztyn et al. (2018) provide field-experimental evidence on status goods (credit cards from an Indonesian bank), which could be interpreted as a manifestation of distinction in our model.[57] Macchi (2023) provides evidence on credit-worthiness signalling strategies in Uganda, showing that obesity facilitates credit access. Interpreting these strategies as aimed at signalling not only wealth (privately observed "head start" in our model) but also reliability (unobserved "talent"), the outcome may correspond to displacement.[58]

6 Alleys for future research

The introduction covered the main insights of the paper. We conclude by briefly evoking three alleys for future research.

Head starts and occupational change. Could the simple model we introduced in this paper be extended to explain occupational change, and particularly, the ongoing polarization of the structure of work in industrialized countries, which features an increasing concentration of employment in high-education, high-wage occupations and low-education, low-wage occupations at the expense of middle-skill occupations (see e.g., Autor 2019)? Such an extension may require introducing multidimensional head starts, and asking which changes in the rel-

[56] Indeed, letting the head start levels of lower-SES, middle-class and higher-SES applicants be given by $0 < m < M$ and $\chi(0) < \chi(m) < \chi(M)$ the associated image values, displacement between the first two and distinction between the last two would obtain whenever

$$\frac{\chi(m) - \chi(0)}{m} < \eta < \frac{\chi(M) - \chi(m)}{M - m}.$$

Alternatively, such differentiated outcomes could be explained by multi-dimensional head starts. As an illustration, suppose head starts are two-dimensional $w = (w_1, w_2)$, with $w_1, w_2 \in \{0, M\}$, such that performance is equal to $y = \theta + e + w_1 + w_2 + \varepsilon$, and that there are different image weights on each head start dimension χ_1, χ_2. Suppose there are three "classes" in the population: the upper class with head start (M, M), the middle class with head start $(0, M)$ and the lower class with head start $(0, 0)$. Then, if $\chi_2 < \eta < \chi_1$, distinction prevails between the upper class and the middle class, while displacement prevails between the middle class and the lower one.

[57] Interestingly, Bursztyn et al. (2018) find that increasing self-esteem causally reduces distinction efforts, which suggests some substitution between social and self images. In our model, the implications of this substitutability can be studied via the agents' intensity of career concerns – e.g., an intervention improving self-esteem would lower μ, which has contrasted consequences depending on whether distinction or displacement prevails (see Section 3.3).

[58] Importantly, Macchi (2023) shows indeed that while obesity is statistically correlated with wealth, it is not interpreted as a signal of beauty nor health, which may suggest that if they could, individuals would otherwise prefer other signals of credit-worthiness.

ative weights of the different head-start dimensions could explain the patterns observed in the data.

Optimal taxation. We focused in this paper on non-monetary head starts, interpreted as either human or social capital. However, head starts may have a monetary component – even if indirect, e.g., the ability to pay for private tutoring or summer camps, or for more comfortable or healthier living conditions. How would optimal policies change given this monetary component? In particular, taxes (or subsidies) may allow the redistribution of part of the head starts across individuals and across generations. Furthermore, whether monetary or not, an agent's head start may be the outcome of the previous generation's effort, which would raise the issue of optimal estate taxation (building on Farhi and Werning 2010).

Markets and morality. Individuals may face both "moral image" concerns and "market image" concerns, the former determined by moral narratives and the latter by production technologies, finite resources and demand and supply equilibria. However, as pointed out by moralists and philosophers, markets and moral narratives may put different weights on each component of "merit" – innate talent, effort and head start. In the wake of works such as Weber's (1905) seminal study on the protestant ethic and the spirit of capitalism, could our model shed some light on the (joint) relations between different production technologies, modes of organization and moral narratives?

References

Alesina, A. and Angeletos, G.-M. (2005). "Fairness and Redistribution". *American Economic Review*, 95(4):960–980.

Ali, S. N. and Bénabou, R. (2020). "Image versus Information: Changing Societal Norms and Optimal Privacy". *American Economic Journal: Microeconomics*, 12(3):116–64.

Andre, P. (2022). "Shallow Meritocracy". R&R at *Review of Economic Studies.*

Arrow, K. J. (1973). "The Theory of Discrimination". In Ashenfelter, O. and Rees, A., editors, *Discrimination in Labor Markets.* Princeton University Press.

Ashraf, N., Bandiera, O., and Lee, S. S. (2014). "Awards Unbundled: Evidence from a Natural Field Experiment". *Journal of Economic Behavior & Organization*, 100:44–63.

Austen-Smith, D. and Fryer, R. G. (2005). "An Economic Analysis of "Acting White"". *Quarterly Journal of Economics*, 120(2):551–583.

Autor, D. H. (2019). "Work of the Past, Work of the Future". *AEA Papers and Proceedings*, 109:1–32.

Baker, G., Gibbons, R., and Murphy, K. J. (1994). "Subjective Performance Measures in Optimal Incentive Contracts". *Quarterly Journal of Economics*, 109(4):1125–1156.

Banks, J. S. and Sobel, J. (1987). "Equilibrium Selection in Signaling Games". *Econometrica*, 55(3):647–661.

Bénabou, R. (1993). "Workings of a City: Location, Education, and Production". *Quarterly Journal of Economics*, 108(3):619–652.

Bénabou, R. and Tirole, J. (2006). "Belief in a Just World and Redistributive Politics". *Quarterly Journal of Economics*, 121(2):699–746.

Bénabou, R. and Tirole, J. (2016). "Bonus Culture: Competitive Pay, Screening, and Multitasking". *Journal of Political Economy*, 124(2):305–370.

Board, S. (2009). "Monopolistic Group Design with Peer Effects". *Theoretical Economics*, 4(1):89–125.

Bonatti, A. and Hörner, J. (2017). "Career Concerns with Exponential Learning". *Theoretical Economics*, 12(1):425–475.

Boudon, R. (1973). *L'inégalité des chances. La mobilité sociale dans les sociétés industrielles.* Armand Colin.

Bourdieu, P. (1979). *La Distinction: Critique sociale du jugement.* Editions de Minuit.

Bourdieu, P. and Passeron, J.-C. (1970). *La Reproduction.* Editions de Minuit.

Bourdieu, P. and Saint-Martin, M. (1970). "L'excellence scolaire et les valeurs du système d'enseignement français". *Annales. Economies, sociétés, civilisations. 25e année*, 1:147–175.

Burdett, K. and Coles, M. G. (1997). "Marriage and Class". *Quarterly Journal of Economics*, 112(1):141–168.

Bursztyn, L., Egorov, G., and Jensen, R. (2018). "Cool to be Smart or Smart to be Cool? Understanding Peer Pressure in Education". *Review of Economic Studies*, 86(4):1487–1526.

Bursztyn, L., Ferman, B., Fiorin, S., Kanz, M., and Rao, G. (2017). "Status Goods: Experimental Evidence from Platinum Credit Cards". *Quarterly Journal of Economics*, 133(3):1561–1595.

Butera, L., Metcalfe, R., Morrison, W., and Taubinsky, D. (2022). "Measuring the Welfare Effects of Shame and Pride". *American Economic Review*, 112(1):122–68.

Cho, I.-K. and Kreps, D. M. (1987). "Signaling Games and Stable Equilibria". *Quarterly Journal of Economics*, 102(2):179–221.

Cisternas, G. (2018). "Career Concerns and the Nature of Skills". *American Economic Journal: Microeconomics*, 10(2):152–89.

Coase, R. (1937). "The Nature of the Firm". *Economica*, 4:386–405.

Coate, S. and Loury, G. C. (1993). "Will Affirmative-Action Policies Eliminate Negative Stereotypes?". *American Economic Review*, 83(5):1220–1240.

Cole, H., Mailath, G., and Postlewaite, A. (1992). "Social Norms, Savings Behavior, and Growth". *Journal of Political Economy*, 100(6):1092–1125.

Dewatripont, M., Jewitt, I., and Tirole, J. (1999a). "The Economics of Career Concerns, Part I: Comparing Information Structures". *Review of Economic Studies*, 66(1):183–198.

Dewatripont, M., Jewitt, I., and Tirole, J. (1999b). "The Economics of Career Concerns, Part II: Application to Missions and Accountability of Government Agencies". *Review of Economic Studies*, 66(1):199–217.

Farhi, E. and Werning, I. (2010). "Progressive Estate Taxation". *Quarterly Journal of Economics*, 125(2):635–673.

Fernández, R. and Galí, J. (1999). "To Each According to ...? Markets, Tournaments, and the Matching Problem with Borrowing Constraints". *Review of Economic Studies*, 66(4):799–824.

Fernández, R. and Rogerson, R. (1995). "On the Political Economy of Education Subsidies". *Review of Economic Studies*, 62(2):249–262.

Frank, R. H. (1985). *Choosing the Right Pond: Human Behavior and the Quest for Status*. Oxford University Press.

Fudenberg, D. and Tirole, J. (1991). "Perfect Bayesian Equilibrium and Sequential Equilibrium". *Journal of Economic Theory*, 53(2):236–260.

Gibbons, R. and Murphy, K. J. (1992). "Optimal Incentive Contracts in the Presence of Career Concerns: Theory and Evidence". *Journal of Political Economy*, 100(3):468–505.

Green, J. R. and Stokey, N. L. (1983). "A Comparison of Tournaments and Contracts". *Journal of Political Economy*, 91(3):349–364.

Holmström, B. (1999). "Managerial Incentive Problems: A Dynamic Perspective". *Review of Economic Studies*, 66(1):169–182.

Langtry, A. (2022). "Keeping Up with "The Joneses": Reference Dependent Choice with Social Comparisons". *American Economic Journal: Microeconomics* (forthcoming).

Macchi, E. (2023). "Worth Your Weight: Experimental Evidence on the Benefits of Obesity in Low-Income Countries". *American Economic Review*, 113(9):2287–2322.

MacLeod, W. B. and Urquiola, M. (2015). "Reputation and School Competition". *American Economic Review*, 105(11):3471–88.

Madsen, E., Williams, B., and Skrzypacz, A. (2022). "Incentive Design for Talent Discovery". mimeo.

Markovits, D. (2019). *The Meritocracy Trap*. Penguin Press.

Merton, R. K. (1957). *Social Theory and Social Structure*. Glencoe, The Free Press.

Moreira, D. and Pérez, S. (2022). "Who Benefits from Meritocracy?". mimeo.

Mulligan, T. (2023). "Meritocracy". In Zalta, E. N. and Nodelman, U., editors, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Fall 2023 edition.

Müller, M. (2022). "Parental Pressure and Intergenerational Mobility in Education". mimeo.

Piketty, T. (1995). "Social Mobility and Redistributive Politics". *Quarterly Journal of Economics*, 110(3):551–584.

Piketty, T. (1998). "Self-Fulfilling Beliefs about Social Status". *Journal of Public Economics*, 70(1):115–132.

Rothschild, C. and Scheuer, F. (2016). "Optimal Taxation with Rent-Seeking". *Review of Economic Studies*, 83(3 (296)):1225–1262.

Sandel, M. J. (2020). *The Tyranny of Merit*. Farrar, Straus and Giroux.

Spence, M. (1973). "Job Market Signaling". *Quarterly Journal of Economics*, 87(3):355–374.

Staab, M. (2022). "The Formation of Social Groups under Status Concern". R&R at *Journal of Economic Theory*.

Weber, M. (2002). *The Protestant Ethic and the Spirit of Capitalism.* Penguin Books. (first publication 1905).

Young, M. (1958). *The Rise of the Meritocracy.* Thames and Hudson.

Appendix

A Sorting and off-path beliefs

Let us explicit our equilibrium refinement in the case of two head start levels ($w \in \{0, M\}$) and degenerate off-path beliefs. Let the audience's off-path beliefs be described by the function $p : h \mapsto p(h)$ where $p(h) \in \{0, 1\}$ is the probability attributed to an agent choosing activity h being rich. Fix an equilibrium, and consider an out-of-equilibrium activity choice h. Denote by $u^*(w)$ the expected equilibrium payoff of an agent with head start w, and by $u(w, p(h), h)$ the expected payoff of an agent with head start w when choosing activity h given the audience's beliefs p.

Definition (**Equilibrium concept, degenerate off-path beliefs**). *We consider Bayesian Perfect equilibria as defined by Fudenberg and Tirole (1991) that further satisfy the following additional requirement: If for an off-path activity h and for a head start level w, there exists another head start level w' such that for any $p(h) \in \{0, 1\}$,*

$$u(w', p(h), h) - u^*(w') \geq u(w, p(h), h) - u^*(w),$$

with a strict inequality for some $p(h)$, then the audience's equilibrium belief should put zero probability on an agent who chooses activity h having head start w.

An important implication of our equilibrium concept is that, with degenerate off-path beliefs, off-path deviations towards activities with lower precision are attributed to poor agents, while off-path deviations towards activities with higher precision are attributed to rich agents. Namely, we have the following result.

Lemma A.1 (**Sorting and degenerate off-path beliefs**). *Suppose there exists a separating equilibrium with degenerate off-path beliefs in which a strictly positive mass of poor agents chooses an activity with precision h_P, while a strictly positive mass of rich agents chooses an activity with precision h_R. Then, $h_P < h_R$. Moreover, any off-path deviation to an activity with precision $h < h_P$ is attributed to a poor agent with probability 1, and any off-path deviation to an activity with precision $h > h_R$ is attributed to a rich agent with probability 1.*

Proof. Consider an equilibrium as described in the Lemma. Let $p(h) \in \{0, 1\}$ denote the audience's belief that an agent in activity with precision h is rich.

Necessary conditions for such an equilibrium to exist are that a poor agent in activity h_P, resp. a rich agent in activity h_R has no strict incentive to deviate to activity h_R, resp.

h_P. Hence,

$$\begin{cases} U(h_P) - \mu\Big(\frac{h_P}{h_0+h_P} - \chi\Big)p(h_P)M \;\geq\; U(h_R) - \mu\Big(\frac{h_R}{h_0+h_R} - \chi\Big)p(h_R)M, \\ U(h_R) + \frac{\mu h_R}{h_0+h_R}M - \mu\Big(\frac{h_R}{h_0+h_R} - \chi\Big)p(h_R)M \\ \qquad\qquad \geq\; U(h_P) + \frac{\mu h_P}{h_0+h_P}M - \mu\Big(\frac{h_P}{h_0+h_P} - \chi\Big)p(h_P)M, \end{cases} \tag{7}$$

and thus

$$\mu\Big(\frac{h_P}{h_0+h_P} - \frac{h_R}{h_0+h_R}\Big)M \leq 0,$$

i.e. $h_P \leq h_R$.

Consider now an off-path deviation to an activity with precision $h < h_P$. For any belief $p' \in \{0,1\}$ about the probability that an agent choosing activity h is rich, a poor agent's gain from deviating from activity h_P to activity h is equal to

$$U(h) - \mu\Big(\frac{h}{h_0+h} - \chi\Big)p'M - \Big[U(h_P) - \mu\Big(\frac{h_P}{h_0+h_P} - \chi\Big)p(h_P)M\Big]$$

whereas a rich agent's gain from deviating from activity h_R to activity h is equal to

$$\begin{aligned} & U(h) + \frac{\mu h}{h_0+h}M - \mu\Big(\frac{h}{h_0+h} - \chi\Big)p'M - \Big[U(h_R) + \frac{\mu h_R}{h_0+h_R}M - \mu\Big(\frac{h_R}{h_0+h_R} - \chi\Big)p(h_R)M\Big] \\ & \leq U(h) + \frac{\mu h}{h_0+h}M - \mu\Big(\frac{h}{h_0+h} - \chi\Big)p'M - \Big[U(h_P) + \frac{\mu h_P}{h_0+h_P}M - \mu\Big(\frac{h_P}{h_0+h_P} - \chi\Big)p(h_P)M\Big] \\ & < U(h) - \mu\Big(\frac{h}{h_0+h} - \chi\Big)p'M - \Big[U(h_P) - \mu\Big(\frac{h_P}{h_0+h_P} - \chi\Big)p(h_P)M\Big]. \end{aligned}$$

where the first inequality follows from (7) and the second one from $h < h_P$. Therefore, with our equilibrium concept, the equilibrium belief $p(h)$ that an agent choosing the off-path activity $h < h_P$ is rich is equal to zero. Similarly, (7) implies that the equilibrium belief $p(h)$ for any off-path activity $h > h_R$ is equal to 1. □

B Proof of Lemma 1

The proof follows from the arguments in the text. In particular,

$$U'(h) = \Big(\mu\eta - g'(e^*(h))\Big)\frac{de^*(h)}{dh} = \Big(\mu\eta - g'(e^*(h))\Big)\frac{1}{g''(e^*(h))}\frac{h_0}{(h_0+h)^2}.$$

Hence, $\max_h U(h) = U(h^*)$ such that $h^*/(h_0+h^*) = \eta$.

C Proofs of Proposition 1 and Corollary 1

C.1 Proof of Proposition 1

Suppose Assumption 1 holds. Hence, $U(h^*) > U(h)$ for any $h \neq h^*$, and therefore, absent head-start inequality ($M = 0$), the unique equilibrium is all agents choosing activity h^*.

Beliefs. Following on our preliminary remark, our equilibrium concept yields that in any equilibrium in which a strictly positive mass of poor agents choose an activity with precision h_P, any off-path deviation to an activity with precision $h < h_P$ is attributed to a poor agent with probability 1 (see Lemma A.1, Appendix A). Similarly, in any equilibrium in which a strictly positive mass of rich agents choose an activity with precision h_R, any off-path deviation to an activity with precision $h > h_R$ is attributed to a rich agent with probability 1.

Hence, let h_P and h_R be resp. the lowest activity (in terms of precision) chosen by a strictly positive mass of poor agents, and h_R the higest activity chosen by a strictly positive mass of rich agents, and let p_X, $X \in \{P, R\}$ be the belief that an agent in activity h_X is rich. The no-profitable-deviation conditions for poor and rich agents require in particular that[59]

$$\begin{aligned} U(h_P) - \mu\Big(\frac{h_P}{h_0 + h_P} - \chi\Big)p_P M &\geq \max_{h \leq h_P} U(h), \\ U(h_R) + \frac{\mu h_R}{h_0 + h_R}M - \mu\Big(\frac{h_R}{h_0 + h_R} - \chi\Big)p_R M &\geq \max_{h \geq h_R} U(h) + \chi M. \end{aligned}$$

Separating equilibria. By Assumption 1, $U(h)$ strictly increases with $h \in (0, h^*)$ and strictly decreases with $h \in (h^*, +\infty)$. As a consequence, in any separating equilibrium, all poor agents choose the same activity, denoted by h_P, while all rich agents choose the same activity h_R. By Lemma A.1, $h_P < h_R$. Our preliminary remark together with Assumption 1 yield that

$$h_P \leq h^* \qquad \text{and} \qquad h_R \geq h^*. \tag{8}$$

As noted in the text, for the poor and the rich not to be tempted to deviate to the other group's activity, the following condition must hold:

$$\mu\Big(\frac{h_P}{h_0 + h_P} - \chi\Big)M \leq U(h_R) - U(h_P) \leq \mu\Big(\frac{h_R}{h_0 + h_R} - \chi\Big)M. \tag{9}$$

With our equilibrium concept, off-path deviations to an activity $h \in (h_P, h_R)$ are attributed

[59] We use the continuity of all expressions with respect to $h \in (0, +\infty)$.

to poor agents with probability 1 if for any belief $p \in [0,1]$,

$$U(h) - \mu\Big(\frac{h}{h_0+h} - \chi\Big)pM - U(h_P) > U(h) + \frac{\mu h}{h_0+h} - \mu\Big(\frac{h}{h_0+h} - \chi\Big)pM - U(h_R) - \mu\chi M,$$

i.e. if

$$\mu\Big(\frac{h}{h_0+h} - \chi\Big)M < U(h_R) - U(h_P),$$

and to rich agents with probability 1 if

$$\mu\Big(\frac{h}{h_0+h} - \chi\Big)M > U(h_R) - U(h_P).$$

Let h' be such that

$$\mu\Big(\frac{h'}{h_0+h'} - \chi\Big)M = U(h_R) - U(h_P).$$

Then, condition (9) implies that $h' \in [h_P, h_R]$, and the necessary and sufficient existence conditions thus write as

$$U(h_P) = \max_{0 \le h \le h'} U(h) \qquad \text{and} \qquad U(h_R) = \max_{h' \le h \le +\infty} U(h).$$

Therefore, with Assumption 1, two cases arise (that are mutually exclusive as we will see shortly):[60]

(i) $h' > h^*$, and then $h_P = h^*$ and $h_R = h' > h_P$, i.e.

$$\mu\Big(\frac{h_R}{h_0+h_R} - \chi\Big)M = U(h_R) - U(h^*). \tag{10}$$

(ii) $h' < h^*$, and then $h_R = h^*$ and $h_P = h' < h_R$, i.e.

$$\mu\Big(\frac{h_P}{h_0+h_P} - \chi\Big)M = U(h^*) - U(h_P). \tag{11}$$

Lastly, again by Assumption 1, $h_R > h_P = h^*$ implies that $U(h_R) < U(h^*)$, whereas $h_P < h_R = h^*$ implies that $U(h^*) > U(h_P)$. Hence, case (i) corresponds to $h^*/(h_0+h^*) < h_R/(h_0+h_R) < \chi$ (distinction), while case (ii) corresponds to $h^*/(h_0+h^*) > h_P/(h_0+h_R) > \chi$ (displacement). This further establishes that the two cases are mutually exclusive.

Comparative statics of h_P and h_R. The monotonicity of h_P and h_R with respect to M

[60] If $h' = h^*$, then both the poor and the rich choose activity h^*, a contradiction.

obtain with the implicit function theorem by differentiating (11) and (10) as the function U strictly increases with $h \in (0, h^*)$ and strictly decreases with $h \in (h^*, +\infty)$. The same argument yields the monotonicity of h_P and h_R with respect to χ and η (noting for η that $h^* < h_R$ under distinction implies $e^*(h^*) < e^*(h_R)$, while $h_P < h^*$ under displacement implies $e^*(h_P) < e^*(h^*)$).

Remark: Comparative statics with respect to μ. When distinction prevails,

$$\begin{aligned}
&\left(U'(h_R) - \frac{\mu h_0}{(h_0+h_R)^2}\right)\frac{\partial h_R}{\partial \mu} \\
&= \eta[e^*(h^*) - e^*(h_R)] - [\mu\eta - g'(e^*(h_R))]\frac{1}{g''(e^*(h_R)}\frac{h_R}{h_0+h_R} + \left(\frac{h_R}{h_0+h_R} - \chi\right)M \\
&= \frac{1}{\mu}[g(e^*(h^*)) - g(e^*(h_R))] - \frac{1}{\mu}[g'(e^*(h^*)) - g'(e^*(h_R))]\frac{g'(e^*(h_R))}{g''(e^*(h_R)} \\
&> \frac{g'(e^*(h_R))}{\mu g''(e^*(h_R))}\left(g''(e^*(h_R))[e^*(h) - e^*(h_R)] - [g'(e^*(h^*)) - g'(e^*(h_R))]\right)
\end{aligned}$$

by strict convexity of g. Hence, if g' is weakly concave, h_R strictly decreases with μ (as $h_R > h^*$ and thus $U'(h_R) < 0$). Similarly, when displacement prevails,

$$\begin{aligned}
&\left(U'(h_P) + \frac{\mu h_0}{(h_0+h_P)^2}\right)\frac{\partial h_P}{\partial \mu} \\
&= \eta[e^*(h^*) - e^*(h_P)] - [\mu\eta - g'(e^*(h_P))]\frac{1}{g''(e^*(h_P)}\frac{h_P}{h_0+h_P} - \left(\frac{h_P}{h_0+h_P} - \chi\right)M \\
&= \frac{1}{\mu}[g(e^*(h^*)) - g(e^*(h_P))] - \frac{1}{\mu}[g'(e^*(h^*)) - g'(e^*(h_P))]\frac{g'(e^*(h_P))}{g''(e^*(h_P)} \\
&> \frac{g'(e^*(h_P))}{\mu g''(e^*(h_P))}\left(g''(e^*(h_P))[e^*(h) - e^*(h_P)] - [g'(e^*(h^*)) - g'(e^*(h_P))]\right)
\end{aligned}$$

by strict convexity of g. Hence, if g' is weakly concave, h_P strictly increases with μ (as $h_P < h^*$ and thus $U'(h_P) < 0$).

Nonetheless, both the rich's and the poor's effort strictly increase with the intensity of image concerns – both under distinction and under displacement.[61]

[61] To see this, note that $g'(e^*(h^*)) = \mu\eta$, while h_R under distinction and h_P under displacement are respectively given by

$$\mu\eta e^*(h_R) - g(e^*(h_R)) = \mu\eta e^*(h^*) - g(e^*(h^*)) + g'(e^*(h_R)) - \mu\chi M,$$

and

$$\mu\eta e^*(h_P) - g(e^*(h_P)) = \mu\eta e^*(h^*) - g(e^*(h^*)) - g'(e^*(h_P)) + \mu\chi M$$

C.2 Proof of Corollary 1

By Proposition 1, the rich's and the poor's expected payoffs in the unique separating equilibrium with degenerate off-path beliefs are given respectively by

$$U(h_R) + \mu\chi M = U(h^*) + \frac{\mu h_R}{h_0 + h_R} M, \qquad \text{resp.} \qquad U(h^*)$$

if distinction prevails ($h^*/(h_0 + h^*) < \chi$), and by

$$U(h^*) + \mu\chi M, \qquad \text{resp.} \qquad U(h_P) = U(h^*) - \mu\Big(\frac{h_P}{h_0 + h_P} - \chi\Big) M.$$

if displacement does ($h^*/(h_0 + h^*) > \chi$).

Hence, the difference between the rich's and the poor's expected payoffs is equal to

$$\frac{\mu h_R}{h_0 + h_R} M, \qquad \text{resp.} \qquad \frac{\mu h_P}{h_0 + h_P} M$$

if distinction prevails, resp. if displacement does.

Comparative statics with respect to M. The result for distinction obtains as h^* does not depend on M, while h_R strictly increases with M. As for displacement, h_P strictly decreases with M, with $h_P < h^*$, and thus $U(h_P)$ strictly decreases with M. Consequently, $\mu[h_P/(h_0 + h_P)]M = U(h^*) + \mu\chi M - U(h_P)$ strictly increases with M, which yields the result for displacement.

Comparative statics with respect to χ. With distinction ($\chi > \eta$), h_R strictly increases with χ, and thus $\mu h_R/(h_0 + h_R)$ ($< \mu\chi$) and $\mu[h_R/(h_0 + h_R)]/U(h^*)$ strictly increase with χ. With displacement ($\chi < \eta$), h_P strictly increases with χ, and thus $\mu h_P/(h_0 + h_P)$ ($> \mu\chi$) strictly increases with χ.

Comparative statics with respect to η. With distinction, h_R strictly increases with η, and thus $\mu h_R/(h_0 + h_R)$ strictly increases with η. Similarly, with displacement, h_P strictly increases with η, and thus $\mu h_P/(h_0 + h_P)$ strictly increases with η.

Comparative statics with respect to μ. With distinction, the difference between the rich's and the poor's expected payoff is equal to

$$\frac{\mu h_R}{h_0 + h_R} M = g'(e^*(h_R))M,$$

and hence the result obtains by noting that $e^*(h_R)$ strictly increases with μ, as shown in the proof of Proposition 1. Similarly, with displacement, the difference between the rich's and the poor's expected payoff is equal to

$$\frac{\mu h_P}{h_0 + h_P} M = g'(e^*(h_P))M,$$

and $e^*(h_P)$ strictly increases with μ, as shown in the proof of Proposition 1.

C.3 Head-start levels

Let us provide the analogue results of Proposition 1 (characterizing the separating equilibria) in cases with more than two head-start levels. The proofs follow from the same arguments as in the proof of Proposition 1 (see Appendix C.1 above), and are thus omitted for the sake of brevity. In particular, the sorting condition remains the same as in the case of two head-start levels.

To alleviate the notation, let $H \equiv h/(h_0 + h)$ for any $h \in \mathbb{R}_+$ and let us index activities by H. In particular, note that with an abuse of notation, $U(\cdot)$ can be written as a function of H, and let accordingly $H^* \equiv h^*/(h_0 + h^*) = \eta$.

Finite number of head-start levels. Suppose that the support of the head starts distribution is the set $\{w_0, ..., w_N\}$, with $N \geq 2$ and $0 = w_0 < ... < w_N = M$. Then, the unique separating equilibrium is:

(i) **(Distinction)** If $\eta < \chi$, the separating equilibrium in which for any $w \in \{w_0, ..., w_N\}$, w-rich agents choose activity $H(w) \in [\eta, \chi)$ such that

$$\begin{cases} H(w_0) = H^* = \eta, \\ U(H(w_{n+1})) - \mu\Big(H(w_{n+1} - \chi\Big)(w_{n+1}) - w_n) = U(H(w_n)) \qquad \text{for } n \in \{0, ..., N-1\}. \end{cases}$$

For any $n, k \in \{0, ..., N\}$, $H(w_k)$ strictly increases with w_n if $k \geq n+1$ and does not depend on w_n if $k \leq n-1$.

(ii) **(Displacement)** If $\eta > \chi$, the separating equilibrium in which for any $w \in \{w_0, ..., w_N\}$, w-rich agents choose activity $H(w) \in [\eta, \chi)$ such that

$$\begin{cases} H(w_N) = H^* = \eta, \\ U(H(w_{n-1})) + \mu\Big(H(w_{n-1} - \chi\Big)(w_n) - w_{n-1}) = U(H(w_n)) \qquad \text{for } n \in \{1, ..., N\}. \end{cases}$$

For any $n, k \in \{0, ..., N\}$, $H(w_k)$ strictly decreases with w_n if $k \leq n-1$ and does not depend on w_n if $k \geq n+1$.

Generalizing the analysis to any finite number of head-start levels delivers an additional observation:

Observation 2. *The larger the number of intermediate head-start levels, the wider apart the extreme precisions.*

Continuum of head-start levels. Suppose that head starts are distributed over $\{0\} \cup [\underline{w}, \overline{w}] \cup \{M\}$ according to some cdf F, strictly increasing over $(\underline{w}, \overline{w}) \subset (0, M)$ and such that $0 < F(0) < F(\overline{w}) < 1$. (The "isolated" boundaries 0 and M conveniently rule out discontinuity issues, as will be clear shortly.) In addition, suppose that the effort cost is quadratic: $g(e) = e^2/2$.

Then, the unique separating equilibrium is

(i) **(Distinction)** If $\eta < \chi$, the separating equilibrium in which w-rich agents choose activity $H(w) \in [\eta, \chi)$ such that

$$\begin{cases} H(0) = H^* = \eta, \\ U(H(\underline{w})) - \mu\big(H(\underline{w}) - \chi\big)\underline{w} = U(H^*), \\ H'(w) = \dfrac{\chi - H(w)}{H(w) - \eta} & \text{for all } w \in (\underline{w}, \overline{w}), \\ U(H(M)) - \mu\big(H(M) - \chi\big)(M - \overline{w}) = U(H(\overline{w})). \end{cases}$$

$H(\cdot)$ is strictly increasing and strictly concave.

(ii) **(Displacement)** If $\eta > \chi$, the separating equilibrium in which w-rich agent choose activity $H(w) \in (\chi, \eta]$ such that

$$\begin{cases} H(M) = H^* = \eta, \\ U(H(\overline{w})) + \mu\big(H(\overline{w}) - \chi\big)(M - \overline{w}) = U(H^*), \\ H'(w) = \dfrac{H(w) - \chi}{\eta - H(w)} & \text{for all } w \in (\underline{w}, \overline{w}), \\ U(H(0)) + \mu\big(H(\underline{w}) - \chi\big)\underline{w} = U(H(\underline{w})). \end{cases}$$

$H(\cdot)$ is strictly increasing and strictly convex.

D Proofs for Section 2.3

D.1 Proof of Lemma 2

Let us first show Lemma 2. By Proposition 1, whenever $\chi \neq \eta$, there exists a unique separating equilibrium in which the rich and the poor choose different activities, with respective precisions $h_R > h_P$, in which they exert efforts $e^*(h_R) > e^*(h_P)$ respectively. Consequently, we turn to (candidate) pooling equilibria, referring to Section 4.2 and Appendix I for additional details.

Lemma D.1. *Let $p \neq 1/2$. In a (candidate) equilibrium in pure strategies in which all agents choose activity $h > 0$, the rich and the poor exert the same effort level $(g')^{-1}(\mu\eta)$ if and only if $\eta = \chi = h/(h_0 + h)$.*

In any candidate equilibrium in pure strategies in activity $h = 0$, the rich and the poor both exert zero effort (see Appendix I), which rules out efficiency.

Proof. Let us consider a (candidate) equilibrium in which both the rich and the poor choose activity $h > 0$ and exert effort $e^\dagger(M)$ and $e^\dagger(0)$ respectively. Suppose that $e^\dagger(M) = e^\dagger(0) > 0$. Then, $e^\dagger(M)$ and $e^\dagger(0)$ are interior and thus satisfy the first-order conditions (25). As $e^\dagger(M) = e^\dagger(0)$, (25) writes as

$$\begin{cases} g'(e^\dagger(M)) &= \dfrac{\mu h}{h_0+h}\Big[1 + (1-p)h_0\Big(\chi - \dfrac{h}{h_0+h}\Big)M^2 E(M, e^\dagger(M))\Big], \\ g'(e^\dagger(0)) &= \dfrac{\mu h}{h_0+h}\Big[1 + (1-p)h_0\Big(\chi - \dfrac{h}{h_0+h}\Big)M^2 E(0, e^\dagger(0))\Big], \end{cases}$$

As $M > 0$ and $e^\dagger(M) = e^\dagger(0)$, $E(M, e^\dagger(M)) = E(0, e^\dagger(0))$ if and only if $p = 1/2$.

Hence, for $p \neq 1/2$, the above system implies:

$$\chi = \frac{h}{h_0 + h}.$$

and thus

$$g'(e^\dagger(M)) = g'(e^\dagger(0)) = \frac{\mu h}{h_0 + h}.$$

Consequently, if the rich and the poor both exert effort $(g')^{-1}(\mu\eta)$, then necessarily:

$$\eta = \frac{h}{h_0 + h} = \chi.$$

Reciprocally, suppose that $\eta = \chi = h/(h_0 + h)$. Then, the arguments in Appendix I imply that the derivative with respect to effort of an agent's payoff in activity h (i.e. the derivate of (24) with respect to e) is equal to

$$\frac{\mu h}{h_0 + h} - g(e),$$

and does not depend on the agent's head start. Therefore, by concavity, the agent's optimal effort in activity h is given by

$$g'(e) = \frac{\mu h}{h_0 + h} = \mu\eta.$$

□

When $p = 1/2$ (non-generic case), there may exist parameter values such that there exist pooling equilibria in an activity h with $h/(h_0 + h) < \min\{\eta, \chi\}$ or $h/(h_0 + h) > \max\{\eta, \chi\}$, that achieve efficiency despite $\eta \neq \chi$.[62]

Lemma D.2. *Let $p \neq 1/2$. If $\eta \neq \chi$, there exists no equilibrium in pure strategies in which the rich and the poor choose the same strictly positive effort. If $\eta = \chi$, there exists a unique equilibrium in pure strategies: all agents choose activity h such that $h/(h_0+h) = \eta$ and exert effort $e^*(h) = (g')^{-1}(\mu\eta)$.*

Proof. Suppose $\eta \neq \chi$. By Lemma D.1, a necessary condition for the existence of a pooling equilibrium in which the rich and the poor choose the same activity h and the same effort $e^\dagger(M) = e^\dagger(0) > 0$, is that $\chi = h/(h_0 + h)$, in which case the agents exert effort $(g')^{-1}(\mu h/(h_0 + h)) = (g')^{-1}(\mu\chi)$.

By (29) (see Appendix I), the agents' equilibrium payoffs $(v(w))_{w\in\{0,M\}}$ are equal to:

$$\begin{cases} v(M) = -(1 - p^\dagger(h, M))\left[\left(\mu\eta - \dfrac{\mu h}{h_0 + h}\right)[M + e^\dagger(M) - e^\dagger(0)] + (\mu\chi - \mu\eta)M\right] \\ \qquad\qquad\qquad\qquad\qquad\qquad +\mu\chi M + \mu\eta e^\dagger(M) - g(e^\dagger(M)), \\ v(0) = p^\dagger(h, 0)\left[\left(\mu\eta - \dfrac{\mu h}{h_0 + h}\right)[M + e^\dagger(M) - e^\dagger(0)] + (\mu\chi - \mu\eta)M\right] + \mu\eta e^\dagger(0) - g(e^\dagger(0)) \end{cases}$$

[62] Using the arguments in Appendix I, a necessary condition for the existence of such an equilibrium with $h/(h_0 + h) < \min\{\eta, \chi\}$ is that a deviation to activity h^* such that $h^*/(h_0 + h^*) = \eta$ be attributed to a poor agent. Hence, using the first-order conditions for the agents' equilibrium effort levels, a necessary condition is that

$$\int_{-\infty}^{+\infty} \frac{e^{-\frac{1}{2}\frac{h_0 h}{h_0+h}(\theta^2+2\theta M+M^2)}}{1 + e^{-\frac{1}{2}\frac{h_0 h}{h_0+h}(2\theta M+M^2)}}\left[\frac{\frac{h_0 h}{h_0+h}M}{1 + e^{-\frac{1}{2}\frac{h_0 h}{h_0+h}(2\theta M+M^2)}} - \frac{e^{\frac{h_0 h}{h_0+h}M^2} - 1}{1 + e^{-\frac{1}{2}\frac{h_0 h}{h_0+h}(2\theta M-M^2)}}\right] d\theta \leq 0,$$

which requires M to be sufficiently high.

where $p^\dagger(h,w) \in (0,1)$ for $w \in \{0, M\}$, and thus for $\chi = h/(h_0+h)$ and $e^\dagger(M) = e^\dagger(0)$,

$$\begin{cases} v(M) = \mu\chi M + \mu\eta e^\dagger(M) - g(e^\dagger(M)), \\ v(0) = \mu\eta e^\dagger(0) - g(e^\dagger(0)), \end{cases}$$

and by our equilibrium refinement, a deviation to an activity h' is thus attributed to a rich agent with probability 1 if

$$\frac{h'}{h_0+h'} > \chi,$$

i.e. if $h' > h$, and to a poor agent if $h' < h$. Suppose that $\chi < \eta$. Then, a rich agent has a strictly profitable deviation to activity h^* such that $h^*/(h_0+h^*) = \eta$ as

$$\mu\eta(g')^{-1}(\mu\eta) - g((g')^{-1}(\mu\eta)) > \mu\eta(g')^{-1}(\mu\chi) - g((g')^{-1}(\mu\chi)).$$

Similarly, if $\chi > \eta$, then a poor agent has a strictly profitable deviation to activity h^* such that $h^*/(h_0+h^*) = \eta$. Therefore, a necessary existence condition for an equilibrium in which all agents choose activity h and the same strictly positive effort level is that

$$\eta = \chi = h/(h_0+h).$$

The above arguments also yield that it is sufficient, as then agents exert effort $(g')^{-1}(\mu\eta)$ which is the unique maximum of $\mu\eta e - g(e)$ by strict convexity of g.

Lastly, let us show that when $\eta = \chi$, there exists no other equilibrium in pure strategies. Suppose by contradiction that there exists an equilibrium in which agents with head start $w \in \{0, M\}$ choose activity h such that $h/(h_0+h) \neq \eta$. If the equilibrium is separating, i.e. if agents with head start $w' \neq w$ choose an activity $h' \neq h$, then agents with head start w have a strictly profitable deviation to activity h^* such that $h^*/(h_0+h^*) = \eta$ regardless of who the deviation is attributed to, as for any $p' \in \{0,1\}$,

$$\begin{aligned} &p'(\mu\chi - \mu\eta)M + \mu\eta w + \mu\eta e^*(h^*) - g(e^*(h^*)) - \Big(\mu\chi w + \mu\eta e^*(h) - g(e^*(h))\Big) \\ &\quad = \mu\eta e^*(h^*) - g(e^*(h^*)) - \Big(\mu\eta e^*(h) - g(e^*(h))\Big) > 0. \end{aligned}$$

Hence, suppose that the equilibrium involves all agents pooling in activity h. If

$$\left(\mu\eta - \frac{\mu h}{h_0+h}\right)[M + e^\dagger(M) - e^\dagger(0)] > 0,$$

then, rich agents have a strictly profitable deviation to activity h^* such that $h^*/(h_0+h^*) = \eta$, regardless of who the deviation is attributed to, as:

$$\mu\eta e^*(h^*) - g(e^*(h^*)) \\ + (1 - p^\dagger(h, M))\Big(\mu\eta - \frac{\mu h}{h_0 + h}\Big)[M + e^\dagger(M) - e^\dagger(0)] - \mu\eta e^\dagger(M) + g(e^\dagger(M)) \;>\; 0$$

by convexity of g. Similarly, if

$$\Big(\mu\eta - \frac{\mu h}{h_0 + h}\Big)[M + e^\dagger(M) - e^\dagger(0)] < 0,$$

then, poor agents have a strictly profitable deviation to activity h^* such that $h^*/(h_0+h^*) = \eta$, regardless of who the deviation is attributed to, as:

$$\mu\eta e^*(h^*) - g(e^*(h^*)) \\ - p^\dagger(h, 0))\Big(\mu\eta - \frac{\mu h}{h_0 + h}\Big)[M + e^\dagger(M) - e^\dagger(0)] - \mu\eta e^\dagger(0) + g(e^\dagger(0)) \;>\; 0.$$

As for the third case, i.e. if

$$\Big(\mu\eta - \frac{\mu h}{h_0 + h}\Big)[M + e^\dagger(M) - e^\dagger(0)] = 0,$$

then as $h/(h_0 + h) \neq \eta$, Lemma D.1 implies that there exists at least one type of agents w such that $e^\dagger(w) \neq (g')^{-1}(\mu\eta)$. Hence, by strict convexity of g, agents with head start w have a strictly profitable deviation to activity h^*, regardless of who the deviation is attributed to. □

Lemma 2 follows from Lemmas D.1 and D.2.

D.2 Several moral views on merit

In our model, *merit* is defined by the relative weights on talent (normalized to 1), effort (η) and head start (χ). Let us briefly sketch three distinct views on merit:

(a) a responsibility-based definition of merit,

(b) an interpretation of head starts as a privately-observable component of talent,

(c) a face-value performance-based definition of merit.

For simplicity, we assume that the activity precision range is $(0, \overline{h})$ for some $\overline{h} < \infty$. [Our analysis in the case $\mathbb{R}_+$ goes through for any $\overline{h}$ sufficiently high.] In addition, we assume that $p \neq 1/2$, ruling out non-generic outcomes.

(a) The responsibility-based definition. The only meritorious dimensions are those for which the agent is individually responsible, i.e. here effort alone. As a consequence, the corresponding weights on talent, effort and head start are $(0, \eta, 0)$. However, since effort is privately observable, a zero weight on talent implies that the agents have zero incentive to exert effort.

A less extreme view of merit may put a strictly positive weight on talent: normalizing it to 1, responsibility-based merit is then defined by $(1, \eta, 0)$, with $\eta \gg 1$. Adapting the arguments in Proposition 1, it can be shown that the unique separating equilibrium is then the separating equilibrium in which the rich choose activity $\overline{h}$, while the poor are displaced to an activity $h_P < \overline{h}$ given by the same condition as in Proposition 1 (replacing h^* by $\overline{h}$. Therefore, (standard) displacement obtains.

(b) Head starts as a privately-observed component of talent. Suppose that an agent's head start is a dimension of the agent's talent, privately learned by the agent in a previous period – e.g. a student who studied a specific course, while still not observing perfectly their talent θ, may have understood whether they enjoy the course topic, or whether they enjoy studying (if so, they enjoy a head start).

The corresponding definition of merit is thus $(1, \eta, 1)$. Two cases arise. If $\eta < 1$ and head-start inequality M is low, then standard distinction (as described in Proposition 1) obtains. Otherwise, if either $\eta \geq 1$ or head-start inequality is large, there exists no separating equilibrium.

(c) Face-value performance-based definition of merit. The different components of merit are valued with their weights in an agent's performance. (Put differently, the "performance technology" is considered to be morally fair and just.) The corresponding definition of merit is thus $(1, 1, 1)$. Then, there exists no separating equilibrium.

D.3 Inequality aversion

To sketch the implications of inequality aversion, let us add to the above efficiency objective, $\mathbb{E}[ae(w) - g(e(w))]$, a quadratic inequality-aversion term:

$$W^{\dagger} \equiv \mathbb{E}[ae(w) - g(e(w))] - r\Big(\mu\eta e^*(h_R) - g(e^*(h_R)) + \mu\chi M - \mu\eta e^*(h_P) + g(e^*(h_P))\Big)^2$$

where $r > 0$ is the coefficient of inequality aversion. Inequality is thus measured here via the difference between the expected payoffs of agents with different head-start levels. Let us focus on equilibria in pure strategies and assume that whenever a separating equilibrium exists (see Proposition 1), it is selected. For expositional simplicity, we further restrict our attention to "intermediate inequality aversion", i.e. values of r such that strictly positive aggregate effort remains optimal.[63]

Lemma D.3 (**Merit and inequality aversion**). *Let $p \neq 1/2$. With intermediate inequality aversion, objective $W^\dagger$ is maximized in a pure-strategy equilibrium if and only if $\chi < \eta \leq a/\mu$.*

Proof. Lemma D.3 follows from the description of the separating equilibria in Proposition 1 and Corollary 1. In particular, the across-head-starts inequality aversion term (second term on the objective's second line) is thus equal to

$$\begin{cases} -r\Big(\dfrac{\mu h_P}{h_0+h_P}M\Big)^2 & \text{if } \chi < \eta, \\ -r\Big(\dfrac{\mu h_R}{h_0+h_R}M\Big)^2 & \text{if } \chi > \eta, \\ -r\big(\mu\eta\chi\big)^2 & \text{if } \chi = \eta, \end{cases}$$

where h_P, h_R are described in Proposition 1 and increase with χ, and where the case $\chi = \mu$ follows from Lemmas D.1-D.2. Moreover, for $\chi = \eta = a/\mu$, all agents are in activity h^* such that $h^*/(h_0 + h^*) = \eta$, and then efficiency is only a second-order concern, while inequality aversion is a first-order concern. As a consequence, maximizing $W^\dagger$ in a pure-strategy equilibrium amounts to solving

$$\max_{\chi,\eta} \left(p\big[ae^*(h_R) - g(e^*(h_R))\big] + (1-p)\big[ae^*(h_P) - g(e^*(h_P))\big] - r\Big(\frac{\mu h_P}{h_0+h_P}M\Big)^2\right)$$
$$\text{s.t.} \qquad \chi \geq 0, \qquad \eta \geq 0, \qquad \chi \leq \eta,$$

for $r \leq \overline{r} < \infty$ ("intermediate inequality aversion") to ensure that the solution satisfies $e^*(h_R) > 0$, $e^*(h_P) > 0$.

The result then follows via standard optimization arguments. □

D.4 Proof of Lemma 3

Lemma 3 follows from the arguments in the text for the outcomes with spot markets for performance, and from Proposition 1 and Corollary 1 for those with meritocracy.

More generally, as a_{spot} increases from 0 to 1, (i) if distinction prevails, meritocracy induces first higher effort and higher inequality, then lower effort and higher inequality, and lastly lower effort and lower inequality than spot markets; (ii) if displacement prevails,

[63] For higher values of r, inequality aversion is so strong that inducing zero effort is optimal, which can be achieved by setting $\chi = \eta = 0$.

meritocracy induces first higher effort and higher inequality, then higher effort and lower inequality, and lastly lower effort and lower inequality.

E Proof of Proposition 2

To alleviate the notation in the rest of this Section, we index by R the rich's activity ($k(M)$) and by P the poor's ($k(0)$). (We focus throughout the proof on deterministic mechanisms and degenerate off-path beliefs.)

We begin with a remark on regarding aligned and countervailing incentives.

Incentives are aligned if a deviation to the outside option is attributed to a poor agent. Hence, with our equilibrium concept, incentives are aligned only if

$$\beta_R + \mu\eta e^*(h_R) - g(e^*(h_R)) - [\beta_P + \mu\eta e^*(h_P) - g(e^*(h_P))] \;\geq\; \mu\Big(\frac{h_{out}}{h_0 + h_{out}} - \chi\Big)M, \quad (12)$$

and necessarily if the inequality is strict. Conversely, incentives are countervailing only if the opposite weak inequality holds, and necessarily if the opposite strict inequality does. In case of equality, we assume that the principal can choose to whom among the rich or the poor the deviation is attributed – i.e. whether incentives are aligned or countervailing.

Aligned incentives. With aligned incentives, the participation constraints write as

$$\begin{cases} \beta_R + \mu\eta e^*(h_R) - g(e^*(h_R)) \geq U_{out} + \mu\Big(\dfrac{h_{out}}{h_0 + h_{out}} - \chi\Big)M, \\ \beta_P + \mu\eta e^*(h_P) - g(e^*(h_P)) \geq U_{out}, \end{cases}$$

and thus (12) implies that the poor's participation constraint is loose only if the rich's participation constraint is loose. As a consequence, a necessary condition for the optimum conditional on aligned incentives (hence conditional on (12)) is that the poor's participation constraint be binding, i.e. $\beta_P + \mu\eta e^*(h_P) - g(e^*(h_P)) = U_{out}$.

Incentive compatibility requires that

$$\begin{aligned} &\mu\Big(\frac{h_R}{h_0 + h_R} - \chi\Big)M \\ &\qquad \geq \beta_R + \mu\eta e^*(h_R) - g(e^*(h_R)) - [\beta_P - \mu\eta e^*(h_P) - g(e^*(h_P))] \\ &\qquad\qquad \geq \mu\Big(\frac{h_P}{h_0 + h_P} - \chi\Big)M, \end{aligned}$$

Hence, incentive compatibility together with (12) imply that $h_R \geq h_{out}$.

Moreover, if $h_P < h_{out}$ at an optimum, then the rich's participation constraint must be binding, and the rich's incentive compatibility constraint be loose at such optimum. This corresponds to the case in which (12) holds with equality.

As a consequence, if $h^{FB} > h_{out}$, then by strict concavity of the objective, a necessary condition for the optimum conditional on aligned incentives is

$$h^{FB} = h_R \geq h_P \geq h_{out},$$

while if $h_{out} > h^{FB}$, a necessary condition is

$$h_{out} = h_R > h_P = h^{FB}.$$

Specifically, conditional on aligned incentives, standard arguments yield that the optimal activity characteristics are given by:

(i) if $h^{FB} > h_{out}$, i.e. if $a + \mu\eta > \mu h_{out}/(h_0 + h_{out})$,

$$\begin{cases} \dfrac{\mu h_R}{h_0 + h_R} = a + \mu\eta, \\ \dfrac{\mu h_P}{h_0 + h_P} = \max\left(\dfrac{\mu h_{out}}{h_0 + h_{out}},\ a + \mu\eta - (1 - q_R)\dfrac{p}{1-p}M\right) \end{cases} \tag{13}$$

for precisions, and for transfers

$$\begin{cases} \beta_R = g(e^*(h_R)) - \mu\eta e^*(h_R) + \mu\Big(\dfrac{h_P}{h_0 + h_P} - \chi\Big)M + U_{out}, \\ \beta_P = g(e^*(h_P)) - \mu\eta e^*(h_P) + U_{out}. \end{cases}$$

(ii) if $h^{FB} < h_{out}$, i.e. if $a + \mu\eta \leq \mu h_{out}/(h_0 + h_{out})$,

$$\begin{cases} \dfrac{\mu h_R}{h_0 + h_R} = \dfrac{\mu h_{out}}{h_0 + h_{out}}, \\ \dfrac{\mu h_P}{h_0 + h_P} = a + \mu\eta \end{cases}$$

for precisions, and for transfers

$$\begin{cases} \beta_R = g(e^*(h_R)) - \mu\eta e^*(h_R) + \mu\Big(\dfrac{h_{out}}{h_0 + h_{out}} - \chi\Big)M + U_{out}, \\ \beta_P = g(e^*(h_P)) - \mu\eta e^*(h_P) + U_{out}. \end{cases}$$

Countervailing incentives. A necessary condition for incentives to be countervailing is thus that (12) hold with the opposite (weak) inequality, i.e. that

$$\beta_R + \mu\eta e^*(h_R) - g(e^*(h_R)) - [\beta_P + \mu\eta e^*(h_P) - g(e^*(h_P))] \ \leq \ \mu\Big(\frac{h_{out}}{h_0 + h_{out}} - \chi\Big)M. \tag{14}$$

With countervailing incentives, the participation constraints write as

$$\begin{cases} \beta_R + \mu\eta e^*(h_R) - g(e^*(h_R)) \geq U_{out}, \\ \beta_P + \mu\eta e^*(h_P) - g(e^*(h_P)) \geq U_{out} - \mu\Big(\dfrac{h_{out}}{h_0 + h_{out}} - \chi\Big)M, \end{cases}$$

and thus (14) implies that the rich's participation constraint is loose only if the poor's participation constraint is loose. As a consequence, a necessary condition for the optimum conditional on countervailing incentives (hence conditional on (14)) is that the rich's participation constraint be binding, i.e. $\beta_R + \mu\eta e^*(h_R) - g(e^*(h_R)) = U_{out}$.

Incentive compatibility (still) requires that

$$\begin{aligned} &\mu\Big(\frac{h_R}{h_0 + h_R} - \chi\Big)M \\ &\qquad\qquad \geq \beta_R + \mu\eta e^*(h_R) - g(e^*(h_R)) - [\beta_P - \mu\eta e^*(h_P) - g(e^*(h_P))] \\ &\qquad\qquad\qquad\qquad\qquad\qquad \geq \mu\Big(\frac{h_P}{h_0 + h_P} - \chi\Big)M, \end{aligned}$$

Hence, incentive compatibility together with (14) imply that $h_{out} \geq h_P$.

Moreover, if $h_R > h_{out}$ at an optimum, then the poor's participation constraint must be binding, and the poor's incentive compatibility constraint be loose at such optimum. (This again corresponds to the case in which (12) holds with equality.)

As a consequence, if $h_{out} > h^{FB}$, then by strict concavity of the objective, a necessary condition for the optimum conditional on countervailing incentives is

$$h_{out} \geq h_R \geq h_P = h^{FB},$$

while if $h^{FB} > h_{out}$, a necessary condition is

$$h^{FB} = h_R > h_P = h_{out}.$$

Specifically, conditional on countervailing incentives, standard arguments yield that the optimal activity characteristics are given by:

(i) if $h_{out} > h^{FB}$, i.e. if $a + \mu\eta < \mu h_{out}/(h_0 + h_{out})$,

$$\begin{cases} \dfrac{\mu h_R}{h_0 + h_R} = \min\Big(a + \mu\eta + (1 - q_P)\dfrac{1-p}{p}M,\ \dfrac{\mu h_{out}}{h_0 + h_{out}}\Big), \\ \dfrac{\mu h_P}{h_0 + h_P} = a + \mu\eta. \end{cases} \tag{15}$$

for precisions, and for transfers

$$\begin{cases} \beta_R = g(e^*(h_R)) - \mu\eta e^*(h_R) + U_{out}, \\ \beta_P = g(e^*(h_P)) - \mu\eta e^*(h_P) - \mu\Big(\dfrac{h_R}{h_0+h_R} - \chi\Big)M + U_{out}. \end{cases}$$

(ii) if $h^{FB} > h_{out}$, i.e. if $a + \mu\eta \geq \mu h_{out}/(h_0 + h_{out})$,

$$\begin{cases} \dfrac{\mu h_R}{h_0+h_R} = a + \mu\eta, \\ \dfrac{\mu h_P}{h_0+h_P} = \dfrac{\mu h_{out}}{h_0+h_{out}}. \end{cases}$$

for precisions, and for transfers

$$\begin{cases} \beta_R = g(e^*(h_R)) - \mu\eta e^*(h_R) + U_{out}, \\ \beta_P = g(e^*(h_P)) - \mu\eta e^*(h_P) - \mu\Big(\dfrac{h_{out}}{h_0+h_{out}} - \chi\Big)M + U_{out}. \end{cases}$$

General case: Choosing aligned or countervailing incentives. Two cases arises depending on whether h_{out} is higher or lower than h^{FB}.

If $h^{FB} > h_{out}$, the principal compares the optimal activity characteristics conditional on aligned incentives, which yield[64]

$$\begin{aligned} p\big[(a+\mu\eta)e^*(h^{FB}) - g(e^*(h^{FB}))\big] &+ (1-p)\big[(a+\mu\eta)e^*(h_P) - g(e^*(h_P))\big] \\ &- (1-q_R)p\frac{\mu h_P}{h_0+h_P}M + p\mu\chi M \end{aligned}$$

where h_P is given by (13), with those conditional on countervailing incentives, which yield

$$\begin{aligned} p\big[(a+\mu\eta)e^*(h^{FB}) - g(e^*(h^{FB}))\big] &+ (1-p)\big[(a+\mu\eta)e^*(h_{out}) - g(e^*(h_{out}))\big] \\ &+ (1-q_P)(1-p)\mu\Big(\frac{h_{out}}{h_0+h_{out}} - \chi\Big)M + q_R p\mu\chi M. \end{aligned}$$

Hence, letting h_P be given by (13), the principal chooses aligned incentives if and only if

$$\begin{aligned} (1-p)\Big((a+\mu\eta)e^*(h_P) - g(e^*(h_P) - \big[(a+\mu\eta)e^*(h_{out}) - g(e^*(h_{out}))\big]\Big) \\ \geq\ (1-q_R)p\mu\Big(\frac{h_P}{h_0+h_P} - \chi\Big)M + (1-q_P)(1-p)\mu\Big(\frac{h_{out}}{h_0+h_{out}} - \chi\Big)M \end{aligned}$$

[64]We substract the constant term

$$-\big[1 - q_R p - q_P(1-p)\big]U_{out}$$

in the principal's objective, as it does not depend on whether incentives are aligned or countervailing.

and countervailing incentives otherwise.[65] In particular, for M sufficiently high that (13) yields $h_P = h_{out}$, the principal chooses aligned incentives if and only if $h_{out}/(h_0 + h_{out}) < \chi$.

If $h_{out} > h^{FB}$, then the principal similarly compares

$$p\big[(a+\mu\eta)e^*(h_{out}) - g(e^*(h_{out}))\big] + (1-p)\big[(a+\mu\eta)e^*(h^{FB}) - g(e^*(h^{FB}))\big] \\ - (1-q_R)p\frac{\mu h_{out}}{h_0 + h_{out}}M + p\mu\chi M$$

with

$$p\big[(a+\mu\eta)e^*(h_R) - g(e^*(h_R))\big] + (1-p)\big[(a+\mu\eta)e^*(h^{FB}) - g(e^*(h^{FB}))\big] \\ + (1-q_P)(1-p)\mu\Big(\frac{h_R}{h_0 + h_R} - \chi\Big)M + q_R p\mu\chi M$$

where h_R is given by (15). In particular, for M sufficiently high that (15) yields $h_R = h_{out}$, the principal chooses aligned incentives if and only if

$$p\Big((a+\mu\eta)e^*(h_{out}) - g(e^*(h_{out}) - \big[(a+\mu\eta)e^*(h_R) - g(e^*(h_R))\big]\Big) \\ \geq\ (1-q_R)p\mu\Big(\frac{h_{out}}{h_0 + h_{out}} - \chi\Big)M + (1-q_P)(1-p)\mu\Big(\frac{h_R}{h_0 + h_R} - \chi\Big)M,$$

and countervailing incentives otherwise. In particular, for M sufficiently high that (15) yields $h_R = h_{out}$, the principal chooses aligned incentives if and only if $h_{out}/(h_0 + h_{out}) \leq \chi$.

Therefore, for M sufficiently high, the principal chooses aligned incentives if $h_{out}/(h_0 + h_{out}) < \chi$, and countervailing incentives if $h_{out}/(h_0 + h_{out}) \geq \chi$.

E.1 Endogenous outside options

Let us now assume that the agents' outside option is a distant, competitive labor market as described in Section 3.1.

Displacement. Suppose $a + \mu\eta > \mu\chi$. By Proposition 1, the precisions chosen by the rich and the poor, $h_{out,R}$ and $h_{out,P}$, in that market are given by $\mu h_{out,R}/(h_0 + h_{out,R}) = a + \mu\eta$ and

$$U(h_{out,P}) = U(h_{out,R}) - \mu\Big(\frac{h_{out,P}}{h_0 + h_{out,P}} - \chi\Big)M,$$

[65]We assume that the indifference case is resolved in favor of aligned incentives when $h^{FB} > h_{out}$, and in favor of countervailing incentives when $h^{FB} < h_{out}$.

where $U(h_{out,X}) \equiv (a+\mu\eta)e^*(h_{out,X}) - g(e^*(h_{out,X})) - d$, for $X \in \{P, R\}$. Hence, with our previous notation,

$$h_{out,R} = h^{FB} > h_{out,P}.$$

Moreover, deviations to $h_{out,R}$ are attributed to rich agents, while deviations to $h_{out,P}$ are attributed to poor agents. The agents' participation constraints thus write as

$$\begin{cases} \beta_R + \mu\eta e^*(h_R) - g(e^*(h_R)) \geq U_{out,P} + \mu\Big(\dfrac{h_{out,P}}{h_0 + h_{out,P}} - \chi\Big)M, \\ \beta_P + \mu\eta e^*(h_P) - g(e^*(h_P)) \geq U_{out,P}. \end{cases}$$

Hence, if the poor's participation constraint is binding, incentive compatibility requires that $h_R \geq h_{out,P}$, whereas if the rich's participation constraint is binding, it requires that $h_{out,P} \geq h_P$.

Therefore, by strict concavity, the principal chooses $h_R = h^{FB}$ and h_P given by (13) replacing h_{out} by $h_{out,P}$.

Distinction. Suppose $a + \mu\eta < \mu\chi$. By Proposition 1, the precisions chosen by the rich and the poor, $h_{out,R}$ and $h_{out,P}$, in that market are given by $\mu h_{out,P}/(h_0 + h_{out,P}) = a + \mu\eta$ and

$$U(h_{out,R}) = U(h_{out,P}) + \mu\Big(\frac{h_{out,R}}{h_0 + h_{out,R}} - \chi\Big)M,$$

where $U(h_{out,X}) \equiv (a+\mu\eta)e^*(h_{out,X}) - g(e^*(h_{out,X})) - d$, for $X \in \{P, R\}$. Hence, with our previous notation,

$$h_{out,P} = h^{FB} < h_{out,R}.$$

Moreover, and as the displacement case, deviations to $h_{out,R}$ are attributed to rich agents, while deviations to $h_{out,P}$ are attributed to poor agents. The agents' participation constraints thus write as

$$\begin{cases} \beta_R + \mu\eta e^*(h_R) - g(e^*(h_R)) \geq U_{out,R}, \\ \beta_P + \mu\eta e^*(h_P) - g(e^*(h_P)) \geq U_{out,R} - \mu\Big(\dfrac{h_{out,R}}{h_0 + h_{out,R}} - \chi\Big)M. \end{cases}$$

Hence, if the poor's participation constraint is binding, incentive compatibility requires that $h_R \geq h_{out,R}$, whereas if the rich's participation constraint is binding, it requires that $h_{out,R} \geq h_P$.

Therefore, by strict concavity, the principal chooses $h_P = h^{FB}$ and h_R given by (15) replacing h_{out} by $h_{out,R}$.

F Proofs of Proposition 3 and Corollary 2

The argument is analogous to the one for the proof of Proposition 1 (see Appendix C).

Suppose that the principal sets a precision cap $\overline{h} < h_R$ with h_R the activity chosen by the rich under *laissez-faire*.

Distinction. Suppose $h^*/(h_0+h^*) < \chi$. The same argument as in the proof of Proposition 1 (see Appendix C) yields that there exists no separating equilibrium.

Displacement. Suppose $h^*/(h_0 + h^*) > \chi$. If $\overline{h}/(h_0 + \overline{h}) < \chi$, the cap forces the agents into the distinction region, and the previous analysis applies, yielding that there exists no separating equilibrium. Hence, suppose $\overline{h}/(h_0 + \overline{h}) > \chi$. Let us first show that the rich choosing activity $\overline{h}$ and the poor activity $h_P(\overline{h})$ such that

$$U(h_P(\overline{h})) = U(\overline{h}) - \mu\Big(\frac{h_P(\overline{h})}{h_0 + h_P(\overline{h})} - \chi\Big)M,$$

is an equilibrium. Any deviation to activities $h < h_P(\overline{h})$ is most attractive to a poor agent and thus, under D1, attributed to a poor agent. Any such deviation is not profitable for a poor agent (and thus for a rich one) as $U(h) < U(h_P(\overline{h}))$. Similarly, by definition of $h_P(\overline{h})$, any deviation to activities $h \in (h_P(\overline{h}), \overline{h}$ is most profitable to a rich agent and thus, under D1, attributed to a rich agent. Any such deviation is not profitable for a rich agent (and thus for a poor one) as $U(h) < U(\overline{h})$. This establishes existence.

Let us now show uniqueness among separating equilibria in pure strategies (under D1). Consider a candidate equilibrium with the poor in activity h and the rich in activity $h' > h$. If $h' < \overline{h}$, a rich agent has a strictly profitable deviation to activity $\overline{h}$ as $h' < \overline{h} < h^*$. As a consequence, h' is necessarily equal to $\overline{h}$. If $h > \overline{h}$, a rich agent has a strictly profitable deviation to activity h, while if $h < h_P(\overline{h})$, a poor agent has a strictly profitable deviation to any activity $h + \varepsilon < h_P(\overline{h})$ (as by definition of $h_P(\overline{h})$, a deviation to any such activity is attributed to a poor agent under D1). Therefore, h is necessarily equal to $h_P(\overline{h})$. Using in particular the preliminary remark in the proof of Proposition 1 (see Appendix C) that, in separating equilibria, the rich and the poor cannot be both indifferent over two activities, the same arguments further establish uniqueness among equilibria in mixed strategies (under D1).

Equilibrium payoffs. Suppose $\overline{h} \in (h_P, h_c)$, i.e. $\overline{h} \in (h_P, h^*)$. Following the cap and with respect to *laissez-faire*, the impact on the poor's payoff is equal to

$$U(h_P(\overline{h})) - U(h^*) < 0,$$

while the impact on the rich's payoff is equal to

$$U(\overline{h}) - U(h^*) < 0.$$

G Proof of Proposition 4

Lemma 4 follows from Proposition 1, and its proof is thus omitted. The comparative statics of h_R and h_P follow from standard computations.[66]

For any $\tau \leq 1$, let

$$W(\tau, M) \equiv p\Big(\mu\eta e^*(h_R(\tau)) - g(e^*(h_R(\tau)))\Big) + (1-p)\Big(\mu\eta e^*(h_P(\tau)) - g(e^*(h_P(\tau)))\Big)$$

[66]Namely, if distinction prevails ($\eta < \chi$), $e^{\ddagger}_P(\tau)$ is given by $g'(e^{\ddagger}_P(\tau)) = (1-\tau)\mu\eta$, and $h_R(\tau)$ by $(1-\tau)\mu h_R/(h_0+h_R) = g'(e^{\ddagger}_R(\tau))$, where $e^{\ddagger}_R(\tau)$ is given by

$$(1-\tau)\mu\eta e^{\ddagger}_R(\tau) - g(e^{\ddagger}_R(\tau)) = (1-\tau)\mu\eta e^{\ddagger}_P(\tau) - g(e^{\ddagger}_P(\tau)) + \big[g'(e^{\ddagger}_R(\tau)) - (1-\tau)\mu\chi\big]M.$$

Therefore, by differentiation,

$$\begin{aligned}
&\Big[(1-\tau)\mu\eta - g'(e^{\ddagger}_R(\tau)) - g''(e^{\ddagger}_R(\tau))M\Big]\frac{(1-\tau)^2}{g''(e^{\ddagger}_R(\tau))}\frac{h_0}{(h_0+h_R)^2}\frac{\partial h_R(\tau)}{\partial \tau} \\
&= \Big[g'(e^{\ddagger}_P(\tau)) - g'(e^{\ddagger}_R(\tau))\Big]\frac{g'(e^{\ddagger}_R(\tau))}{g''(e^{\ddagger}_R(\tau))} + \big[g(e^{\ddagger}_R(\tau)) - g(e^{\ddagger}_P(\tau))\big] \\
&< -\Big[g'(e^{\ddagger}_R(\tau)) - g'(e^{\ddagger}_P(\tau))\Big]\frac{g'(e^{\ddagger}_R(\tau))}{g''(e^{\ddagger}_R(\tau))} + \big[e^{\ddagger}_R(\tau)) - e^{\ddagger}_P(\tau))\big]g'(e^{\ddagger}_R(\tau))
\end{aligned}$$

where the inequality obtains by convexity of g. Hence, if g' is (weakly) concave, the RHS is strictly negative, and thus $\partial h_R/\partial\tau > 0$ (as the term between brackets on the LHS is strictly negative).

Similarly, if displacement prevails ($\eta > \chi$), $e^{\ddagger}_R(\tau)$ is given by $g'(e^{\ddagger}_R(\tau)) = (1-\tau)\mu\eta$, and $h_P(\tau)$ by $(1-\tau)\mu h_P/(h_0+h_P) = g'(e^{\ddagger}_P(\tau))$, where $e^{\ddagger}_P(\tau)$ is given by

$$(1-\tau)\mu\eta e^{\ddagger}_P(\tau) - g(e^{\ddagger}_P(\tau)) = (1-\tau)\mu\eta e^{\ddagger}_R(\tau) - g(e^{\ddagger}_R(\tau)) - \big[g'(e^{\ddagger}_P(\tau)) - (1-\tau)\mu\chi\big]M.$$

Therefore, by differentiation,

$$\begin{aligned}
&\Big[(1-\tau)\mu\eta - g'(e^{\ddagger}_P(\tau)) + g''(e^{\ddagger}_P(\tau))M\Big]\frac{(1-\tau)^2}{g''(e^{\ddagger}_P(\tau))}\frac{h_0}{(h_0+h_P)^2}\frac{\partial h_P(\tau)}{\partial \tau} \\
&= \Big[g'(e^{\ddagger}_R(\tau)) - g'(e^{\ddagger}_P(\tau))\Big]\frac{g'(e^{\ddagger}_P(\tau))}{g''(e^{\ddagger}_P(\tau))} - \big[g(e^{\ddagger}_R(\tau)) - g(e^{\ddagger}_P(\tau))\big] \\
&< \Big[g'(e^{\ddagger}_R(\tau)) - g'(e^{\ddagger}_P(\tau))\Big]\frac{g'(e^{\ddagger}_P(\tau))}{g''(e^{\ddagger}_P(\tau))} - \big[e^{\ddagger}_R(\tau)) - e^{\ddagger}_P(\tau))\big]g'(e^{\ddagger}_P(\tau))
\end{aligned}$$

where the inequality obtains by convexity of g. Hence, if g' is (weakly) concave, the RHS is strictly negative, and thus $\partial h_R/\partial\tau < 0$ (as the term between brackets on the LHS is strictly positive).

The principal thus solves:

$$\max_{\tau} \; W(\tau, M).$$

Let us define for any $\tau \leq 1$, $e_P^\ddagger : \tau \mapsto e^*(h_P(\tau))$ and $e_R^\ddagger : \tau \mapsto e^*(h_R(\tau))$. Hence, the principal's objective writes as

$$W(\tau, M) = p\Big(\mu\eta e_R^\ddagger(\tau) - g(e_R^\ddagger(\tau))\Big) + (1-p)\Big(\mu\eta e_P^\ddagger(\tau) - g(e_P^\ddagger(\tau))\Big)$$

The objective $W(\tau, M)$ is twice continuously differentiable for $\tau < 1$ and $M > 0$.

Differentation yields that

$$\frac{\partial W}{\partial \tau} = p\Big(\mu\eta - g'(e_R^\ddagger(\tau))\Big)\frac{\partial e_R^\ddagger(\tau)}{\partial \tau} + (1-p)\Big(\mu\eta - g'(e_P^\ddagger(\tau))\Big)\frac{\partial e_P^\ddagger(\tau)}{\partial \tau} \tag{16}$$

and

$$\begin{aligned}\frac{\partial^2 W}{\partial M \partial \tau} = p\Big(\mu\eta - g'(e_R^\ddagger(\tau))\Big)\frac{\partial^2 e_R^\ddagger(\tau)}{\partial M \partial \tau} - pg''(e_R^\ddagger(\tau))\frac{\partial e_R^\ddagger(\tau)}{\partial M}\frac{\partial e_R^\ddagger(\tau)}{\partial \tau} \\ + (1-p)\Big(\mu\eta - g'(e_P^\ddagger(\tau))\Big)\frac{\partial^2 e_P^\ddagger(\tau)}{\partial M \partial \tau} - (1-p)g''(e_P^\ddagger(\tau))\frac{\partial e_P^\ddagger(\tau)}{\partial M}\frac{\partial e_P^\ddagger(\tau)}{\partial \tau}\end{aligned} \tag{17}$$

Distinction. Suppose $\eta < \chi$. By Lemma 4, in the unique separating equilibrium with degenerate off-path beliefs, $e_P^\ddagger(\tau)$ is given by $g'(e_P^\ddagger(\tau)) = (1-\tau)\mu\eta$, and $e_R^\ddagger(\tau)$ is given by

$$\begin{aligned}(1-\tau)\mu\eta e_R^\ddagger(\tau) - g(e_R^\ddagger(\tau)) &= (1-\tau)\mu\eta e_P^\ddagger(\tau) - g(e_P^\ddagger(\tau)) + (1-\tau)\mu\Big(\frac{h_R}{h_0 + h_R} - \chi\Big)M \\ &= (1-\tau)\mu\eta e_P^\ddagger(\tau) - g(e_P^\ddagger(\tau)) + [g'(e_R^\ddagger(\tau)) - (1-\tau)\mu\chi]M.\end{aligned}$$

Therefore,

$$\frac{\partial e_P^\ddagger(\tau)}{\partial \tau} = -\mu\eta, \qquad \text{and} \qquad \frac{\partial e_P^\ddagger(\tau)}{\partial M} = 0,$$

while

$$\Big[(1-\tau)\mu\eta - g'(e_R^\ddagger(\tau)) - g''(e_R^\ddagger(\tau))M\Big]\frac{\partial e_R^\ddagger(\tau)}{\partial \tau} = \mu\eta[e_R^\ddagger(\tau) - e_P^\ddagger(\tau)] + \mu\chi M > 0$$

where the inequality follows from $e_R^\ddagger(\tau) > e_P^\ddagger(\tau)$ for any $\tau < 1$ (which further implies that the term between brackets on the LHS is strictly negative), and

$$\Big[(1-\tau)\mu\eta - g'(e_R^\ddagger(\tau)) - g''(e_R^\ddagger(\tau))M\Big]\frac{\partial e_R^\ddagger(\tau)}{\partial M} = g'(e_R^\ddagger(\tau)) - (1-\tau)\mu\chi < 0,$$

and thus

$$\begin{aligned}\Big[(1-\tau)\mu\eta - g'(e_R^\ddagger(\tau)) - g''(e_R^\ddagger(\tau))M\Big]&\frac{\partial^2 e_R^\ddagger(\tau)}{\partial M\partial\tau} \\ = \Big[g''(e_R^\ddagger(\tau)) + g^{(3)}(e_R^\ddagger(\tau))M\Big]&\frac{\partial e_R^\ddagger(\tau)}{\partial M}\frac{\partial e_R^\ddagger(\tau)}{\partial \tau} + \mu\eta\frac{\partial e_R^\ddagger(\tau)}{\partial M} + \mu\chi\end{aligned} \tag{18}$$

Hence, in particular, (16) writes as

$$\frac{\partial W}{\partial\tau} = p\Big(\mu\eta - g'(e_R^\ddagger(\tau))\Big)\frac{\mu\eta\big[e_R^\ddagger(\tau) - e_P^\ddagger(\tau)\big] + \mu\chi M}{(1-\tau)\mu\eta - g'(e_R^\ddagger(\tau)) - g''(e_R^\ddagger(\tau))M} - (1-p)\tau(\mu\eta)^2 \tag{19}$$

As noted above, by Lemma 4, $\mu\eta < \mu h_R/(h_0 + h_R)$, and therefore, $\partial W/\partial\tau > 0$ for any $\tau \leq 0$. Hence, from (19), any solution τ^* to $\partial W(\tau)/\partial\tau = 0$ is necessarily strictly positive, $\tau^* > 0$, and such that $\mu\eta < g'(e_R^\ddagger(\tau)) = (1-\tau^*)\mu h_R/(h_0 + h_R)$.

In addition, the above computations imply that, if $g''(0) > 0$ (e.g., if g is quadratic and thus $g''(0) = 1$),

$$\lim_{\tau\to 1}\frac{\partial W}{\partial\tau} = -p\frac{\mu\eta\chi}{g''(0)} - (1-p)(\mu\eta)^2 < 0.$$

As a consequence, any solution to the principal's program necessarily satisfies the first-order condition $\partial W(\tau)/\partial\tau = 0$ and lies in $(0,1)$. Moreover, as for $\tau = 0$, $\partial W(0)/\partial\tau > 0$, and as W is twice continuously differentiable with respect to τ, there exists such a solution $\tau^* \in (0,1)$, and for any τ in a neighbourhood of τ^*, $\partial^2 W(\tau)/\partial\tau^2 \leq 0$.

Equation (17) writes as

$$\begin{aligned}
\frac{\partial^2 W}{\partial M \partial \tau} &= p\Big(\mu\eta - g'(e_R^{\ddagger}(\tau))\Big)\frac{\partial^2 e_R^{\ddagger}(\tau)}{\partial M \partial \tau} - pg''(e_R^{\ddagger}(\tau))\frac{\partial e_R^{\ddagger}(\tau)}{\partial M}\frac{\partial e_R^{\ddagger}(\tau)}{\partial \tau} \\
&= p\left[\frac{(\mu\eta - g'(e_R^{\ddagger}(\tau)))\left[g''(e_R^{\ddagger}(\tau)) + g^{(3)}(e_R^{\ddagger}(\tau))M\right]}{(1-\tau)\mu\eta - g'(e_R^{\ddagger}(\tau)) - g''(e_R^{\ddagger}(\tau))M} - g''(e_R^{\ddagger}(\tau))\right]\frac{\partial e_R^{\ddagger}(\tau)}{\partial M}\frac{\partial e_R^{\ddagger}(\tau)}{\partial \tau} \\
&\quad + p\frac{\mu\eta - g'(e_R^{\ddagger}(\tau))}{(1-\tau)\mu\eta - g'(e_R^{\ddagger}(\tau)) - g''(e_R^{\ddagger}(\tau))M}\Big(\mu\eta\frac{\partial e_R^{\ddagger}(\tau)}{\partial M} + \mu\chi\Big) \\
&= p\Bigg[\frac{\left[\tau\mu\eta + g''(e_R^{\ddagger}(\tau))M\right]g''(e_R^{\ddagger}(\tau))}{(1-\tau)\mu\eta - g'(e_R^{\ddagger}(\tau)) - g''(e_R^{\ddagger}(\tau))M} \\
&\qquad\qquad + \frac{(\mu\eta - g'(e_R^{\ddagger}(\tau)))g^{(3)}(e_R^{\ddagger}(\tau))M}{(1-\tau)\mu\eta - g'(e_R^{\ddagger}(\tau)) - g''(e_R^{\ddagger}(\tau))M}\Bigg]\frac{\partial e_R^{\ddagger}(\tau)}{\partial M}\frac{\partial e_R^{\ddagger}(\tau)}{\partial \tau} \\
&\quad + p\frac{\mu\eta - g'(e_R^{\ddagger}(\tau))}{(1-\tau)\mu\eta - g'(e_R^{\ddagger}(\tau)) - g''(e_R^{\ddagger}(\tau))M}\Big(\mu\eta\frac{\partial e_R^{\ddagger}(\tau)}{\partial M} + \mu\chi\Big)
\end{aligned}$$

The first term within brackets (first line) is strictly negative and, if $g^{(3)}(e_R^{\ddagger}(\tau)) \leq 0$, the second one (second line) is weakly negative. Moreover, with the above computations, $\partial e_R^{\ddagger}/\partial M > 0$ and $\partial e_R^{\ddagger}/\partial \tau < 0$. Lastly, the term on the third line is strictly positive. Therefore,

$$\frac{\partial^2 W}{\partial M \partial \tau} > 0,$$

which yields the result.

Displacement. Suppose $\eta > \chi$. By Lemma 4, in the unique separating equilibrium with degenerate off-path beliefs, $e_R^{\ddagger}(\tau)$ is given by $g'(e_R^{\ddagger}(\tau)) = (1-\tau)\mu\eta$, and $e_P^{\ddagger}(\tau)$ is given by

$$\begin{aligned}
(1-\tau)\mu\eta e_P^{\ddagger}(\tau) - g(e_P^{\ddagger}(\tau)) &= (1-\tau)\mu\eta e_R^{\ddagger}(\tau) - g(e_R^{\ddagger}(\tau)) - (1-\tau)\mu\Big(\frac{h_P}{h_0 + h_P} - \chi\Big)M \\
&= (1-\tau)\mu\eta e_R^{\ddagger}(\tau) - g(e_R^{\ddagger}(\tau)) - \left[g'(e_P^{\ddagger}(\tau)) - (1-\tau)\mu\chi\right]M.
\end{aligned}$$

Therefore,

$$\frac{\partial e_R^{\ddagger}(\tau)}{\partial \tau} = -\mu\eta, \qquad \text{and} \qquad \frac{\partial e_R^{\ddagger}(\tau)}{\partial M} = 0,$$

while

$$\Big[(1-\tau)\mu\eta - g'(e_P^{\ddagger}(\tau)) + g''(e_P^{\ddagger}(\tau))M\Big]\frac{\partial e_P^{\ddagger}(\tau)}{\partial \tau} = \mu\eta\left[e_P^{\ddagger}(\tau) - e_R^{\ddagger}(\tau)\right] - \mu\chi M < 0$$

where the inequality follows from $e_P^\ddagger(\tau) < e_R^\ddagger(\tau)$ for any $\tau < 1$ (which further implies that the term between brackets on the LHS is strictly negative), and

$$\left[(1-\tau)\mu\eta - g'(e_P^\ddagger(\tau)) + g''(e_P^\ddagger(\tau))M\right]\frac{\partial e_P^\ddagger(\tau)}{\partial M} = -g'(e_P^\ddagger(\tau)) + (1-\tau)\mu\chi > 0,$$

and thus

$$\begin{aligned}&\left[(1-\tau)\mu\eta - g'(e_P^\ddagger(\tau)) + g''(e_P^\ddagger(\tau))M\right]\frac{\partial^2 e_P^\ddagger(\tau)}{\partial M \partial \tau} \\ &\quad = \left[g''(e_P^\ddagger(\tau)) - g^{(3)}(e_P^\ddagger(\tau))M\right]\frac{\partial e_P^\ddagger(\tau)}{\partial M}\frac{\partial e_P^\ddagger(\tau)}{\partial \tau} + \mu\eta\frac{\partial e_P^\ddagger(\tau)}{\partial M} - \mu\chi\end{aligned} \tag{20}$$

Hence, in particular, (16) writes as

$$\frac{\partial W}{\partial \tau} = -p\tau(\mu\eta)^2 + (1-p)\Big(\mu\eta - g'(e_P^\ddagger(\tau))\Big)\frac{\mu\eta\left[e_P^\ddagger(\tau) - e_R^\ddagger(\tau)\right] - \mu\chi M}{(1-\tau)\mu\eta - g'(e_P^\ddagger(\tau)) + g''(e_P^\ddagger(\tau))M} \tag{21}$$

As noted above, by Lemma 4, $\mu\eta > \mu h_P/(h_0+h_P)$, and therefore, $\partial W/\partial\tau < 0$ for any $\tau \geq 0$. Hence, from (19), any solution τ^* to $\partial W(\tau)/\partial\tau = 0$ is necessarily strictly negative, $\tau^* < 0$, and such that $\mu\eta > g'(e_P^\ddagger(\tau)) = (1-\tau^*)\mu h_P/(h_0+h_P)$.

In addition, the above computations imply that for any $\tau < \tau'$ where $\tau' < 0$ is given by $g'(e_P^\ddagger(\tau')) = \mu\eta$,

$$\frac{\partial W}{\partial \tau} > 0.$$

As a consequence, any solution to the principal's program necessarily satisfies the first-order condition $\partial W(\tau)/\partial\tau = 0$ and lies in $(-\infty, 0)$.

As W is twice continuously differentiable with respect to τ, with $\partial W(\tau)/\partial\tau > 0$ for any τ below a strictly negative threshold, and $\partial W(0) < 0$, such a solution τ^* exists, and for any τ in a neighbourhood of τ^*, $\partial^2 W(\tau)/\partial\tau^2 \leq 0$.

Equation (17) writes as

$$\begin{aligned}
\frac{\partial^2 W}{\partial M \partial \tau} &= (1-p)\Big(\mu\eta - g'(e_P^{\ddagger}(\tau))\Big)\frac{\partial^2 e_P^{\ddagger}(\tau)}{\partial M \partial \tau} - (1-p)g''(e_P^{\ddagger}(\tau))\frac{\partial e_P^{\ddagger}(\tau)}{\partial M}\frac{\partial e_P^{\ddagger}(\tau)}{\partial \tau} \\
&= (1-p)\left[\frac{\big(\mu\eta - g'(e_P^{\ddagger}(\tau))\big)\big[g''(e_P^{\ddagger}(\tau)) - g^{(3)}(e_P^{\ddagger}(\tau))M\big]}{(1-\tau)\mu\eta - g'(e_P^{\ddagger}(\tau)) + g''(e_P^{\ddagger}(\tau))M} - g''(e_P^{\ddagger}(\tau))\right]\frac{\partial e_P^{\ddagger}(\tau)}{\partial M}\frac{\partial e_P^{\ddagger}(\tau)}{\partial \tau} \\
&\quad + (1-p)\frac{\mu\eta - g'(e_P^{\ddagger}(\tau))}{(1-\tau)\mu\eta - g'(e_P^{\ddagger}(\tau)) + g''(e_P^{\ddagger}(\tau))M}\left(\mu\eta\frac{\partial e_P^{\ddagger}(\tau)}{\partial M} - \mu\chi\right) \\
&= (1-p)\left[\frac{\big[\tau\mu\eta - g''(e_P^{\ddagger}(\tau))M\big]g''(e_P^{\ddagger}(\tau))}{(1-\tau)\mu\eta - g'(e_P^{\ddagger}(\tau)) + g''(e_P^{\ddagger}(\tau))M}\right. \\
&\qquad \left. - \frac{(\mu\eta - g'(e_P^{\ddagger}(\tau))g^{(3)}(e_P^{\ddagger}(\tau))M}{(1-\tau)\mu\eta - g'(e_P^{\ddagger}(\tau)) + g''(e_P^{\ddagger}(\tau))M}\right]\frac{\partial e_P^{\ddagger}(\tau)}{\partial M}\frac{\partial e_P^{\ddagger}(\tau)}{\partial \tau} \\
&\quad + (1-p)\frac{\mu\eta - g'(e_P^{\ddagger}(\tau))}{(1-\tau)\mu\eta - g'(e_P^{\ddagger}(\tau)) + g''(e_P^{\ddagger}(\tau))M}\left(\mu\eta\frac{\partial e_P^{\ddagger}(\tau)}{\partial M} - \mu\chi\right)
\end{aligned}$$

The first term within brackets (first line) is strictly negative for $\tau < 0$ and, if $g^{(3)}(e_P^{\ddagger}(\tau)) \geq 0$, the second one (second line) is weakly negative. Moreover, with the above computations, $\partial e_R^{\ddagger}/\partial M < 0$ and $\partial e_R^{\ddagger}/\partial \tau < 0$. Lastly, the term on the third line is strictly negative. Therefore,

$$\frac{\partial^2 W}{\partial M \partial \tau} < 0,$$

which yields the result.

H Proof of Proposition 5

Proposition 5 follows directly from the proofs of Lemma A.1 and Proposition 1, as Assumption 1 ensures that $U(h)$ has the required properties for the arguments to go through.

In particular, from the proof of Proposition 1 (see Appendix C), the dominance of either displacement or distinction depends on the sign of the following difference:

$$\frac{\mu h^*}{h_0 + h^*}M - \big[\mu\chi M + c(M)\big],$$

which yields the results with respect to the shape of c. In particular, if interior, the cutoff M^* is given by: $\mu h^*/(h_0 + h^*) = \mu\chi + c(M^*)/M^*$.

In Section H.1 below, we provide details on a result mentioned in the text.

H.1 Non-additively separable transfers

Let us consider transfers non-additively separable in expected effort and head start. We microfound such transfers as stemming from competitive wage-bidding by profit-maximizing firms, or equivalently from competitive fee-setting by profit-maximizing private universities, in a market with free entry. We show that distinction and displacement still happen in such environments.

Namely, let us consider a continuum of organizations vying to attract agents. Each organization chooses the precision h of the activity it requires the agents to perform, and the associated wage/fee β. Each organization that successfully attracts some agents makes a profit $\pi(\mathbb{E}[y])$ per recruited agent, where the argument of $\pi(\cdot)$ is the expected outcome of the organization's members (e.g., firm's employees). We assume that π is positive, strictly increasing and continuously differentiable. Hence, such organizations may be firms whose business involves "collective" tasks, so that the aggregate performance matters for firm performance, or with slightly different conditions, universities interested in the aggregate absolute image of their students (with objective $\pi(\mathbb{E}[\hat{\theta} + \eta\hat{e} + \chi\hat{m}])$). We rely on the "firm" interpretation henceforth.

Timing is as follows: (1) Firms simultaneously commit to a precision h and a wage β; (2) Agents observe the firms' offers and choose which firm to work for. Firms maximize their profits, and face no entry costs nor capacity constraints. Whenever two firms offer the same precision and wage, we assume that agents choose randomly between the two.

Free entry and competition among organizations implies that in equilibrium, each firm offers a wage $\beta = \pi(\mathbb{E}[y])$ to its potential hires, where the expectation depends on the firm's chosen precision and equilibrium beliefs about the agents it will attract.

We assume that $\pi(\cdot)$ is strictly concave.[67] Hence, we define precision h_R^* as the precision that maximizes $[U(h) + \pi(e^*(h) + M)]$, precision h_P^* as the precision that maximizes $[U(h) + \pi(e^*(h))]$, and precision h_a^* as the precision that maximizes $[U(h) + \pi(e^* + pM)]$. Hence, by strict supermodularity, for any $pM > 0$, $h_R^* < h_a^* < h_P^*$.

Proposition H.1 (**Endogenous transfers: Competitive equilibrium with free entry**). *For a given total head start pM in the economy, absent head-start inequality (i.e. redistributing the total head start pM equally across agents), the unique equilibrium is all agents choosing activity $h_a^* \in (h_R^*, h_P^*)$.*

By contrast, with head-start inequality ($w \in \{0, M\}$), there exists $\chi^\dagger < h_P^*/(h_0 + h_P^*)$ and $\chi^\ddagger < h_R^*/(h_0 + h_R^*)$, with $\chi^\dagger > \chi^\ddagger$, such that a separating equilibrium in pure strategies exists only if

[67] The case of linear $\pi(\mathbb{E}[y]) = \varrho\mathbb{E}[y]$ is equivalent to our previous case, changing the weight on expected head start from $\mu\chi$ to $(\mu\chi + \varrho)$.

(i) ***(Distinction)*** $\chi > \chi^{\dagger}$ *and image concerns are sufficiently intense (μ high), in which case the poor choose activity h_P^* and the rich choose an activity $h_R > h_P^*$.*

(ii) ***(Displacement)*** $\chi < \chi^{\ddagger}$ *and image concerns are sufficiently intense (μ high), in which case the rich choose activity h_R^* and the poor choose an activity $h_P < h_R^*$*

In other words, competitive wage formation generates an additional incentive for the poor to pool with the rich, and for the rich to distinguish themselves from the poor. As a consequence and in contrast to exogenous wages, separating equilibria do not exist for some parameters regions with strictly positive measure.

Proof. Let us prove the following result, from which Proposition H.1 immediately follows.

Proposition H.2. *For a given total head start in the economy pM, absent head-start inequality (i.e. redistributing the total head start pM equally across agents), the unique equilibrium is all agents choosing activity $h_a^* \in (h_R^*, h_P^*)$. By contrast, with head-start inequality ($w \in \{0, M\}$, $M > 0$), a separating equilibrium in pure strategies with degenerate off-path beliefs exists if and only if either*

(i) ***(Distinction)*** *The following inequality holds:*

$$\mu\Big(\frac{h_P^*}{h_0 + h_p^*} - \chi\Big)M < \pi(e^*(h_P^*) + M) - \pi(e^*(h_P^*)),$$

and there exists $h_R > h_P^*$ such that

$$U(h_R) + \pi(e^*(h_R) + M) - \mu\Big(\frac{h_R}{h_0 + h_R} - \chi\Big)M = U(h_P^*) + \pi(e^*(h_P^*)),$$

in which case the poor choose activity h_P^* and the rich choose activity h_R.

(ii) ***(Displacement)*** *The following inequality holds:*

$$\mu\Big(\frac{h_R^*}{h_0 + h_R^*} - \chi\Big)M > \pi(e^*(h_R^*) + M) - \pi(e^*(h_R^*)),$$

and there exists $h_P < h_R^*$ such that

$$U(h_P) + \pi(e^*(h_P)) + \mu\Big(\frac{h_P}{h_0 + h_P} - \chi\Big)M = U(h_R^*) + \pi(e^*(h_R^*) + M),$$

in which case the rich choose activity h_R^* and the poor choose activity h_P.

Let $p(h) \in \{0, 1\}$ be the audience's (degenerate) belief that an agent in activity $h \geq 0$ is rich. Free entry and wage competition yield that firms choosing precision h (if any) offer a wage $\pi(e^*(h) + p(h)M)$. The argument then mimicks the one of the proof of Proposition

1 (see Appendix C), replacing the fixed transfer β by a wage $\pi\big(e^*(h) + p(h)M\big)$ which is now a function of precision h and beliefs $p(h)$. For clarity, to single out wages, we denote $U(h) \equiv \mu\eta e^*(h) - g(e^*(h))$.

(Degenerate) Beliefs. Our equilibrium concept thus yields that in any separating equilibrium in which a strictly positive mass of poor agents choose an activity with precision h_P, any off-path deviation to an activity with precision $h < h_P$ is attributed to a poor agent with probability 1 (see Lemma A.1, Appendix A). Similarly, in any equilibrium in which a strictly positive mass of rich agents choose an activity with precision h_R, any off-path deviation to an activity with precision $h > h_R$ is attributed to a rich agent with probability 1.

Hence, consider a separating equilibrium and let h_P and h_R be resp. the highest activity (in terms of precision) chosen by a strictly positive mass of poor agents, and h_R the lowest activity chosen by a strictly positive mass of rich agents. The no-profitable-deviation conditions conditions for poor and rich agents require in particular that[68]

$$\begin{cases} U(h_P) + \pi\big(e^*(h_P) + p(h_P)M\big) - \mu\Big(\dfrac{h_P}{h_0 + h_P} - \chi\Big)p_P M \ \geq \ \max\limits_{h \leq h_P} U(h) + \pi\big(e^*(h)\big), \\ U(h_R) + \pi\big(e^*(h_R) + p(h_R)M\big) + \dfrac{\mu h_R}{h_0 + h_R}M - \mu\Big(\dfrac{h_R}{h_0 + h_R} - \chi\Big)p_R M \\ \qquad\qquad\qquad\qquad\qquad\qquad \geq \ \max\limits_{h \geq h_R} U(h) + \pi\big(e^*(h) + M\big) + \chi M. \end{cases}$$

Separating equilibria. By assumption, $U(h) + \pi(e^*(h) + M)$ strictly increases with $h \in (0, h_R^*)$ and strictly decreases with $h \in (h_R^*, +\infty)$, while $U(h) + \pi(e^*(h))$ strictly increases with $h \in (0, h_P^*)$ and strictly decreases with $h \in (h_P^*, +\infty)$.

In any separating equilibrium in pure strategies, all poor agents choose the same activity, denoted by h_P, while all rich agents choose the same activity h_R. By the same arguments as in the proof of Lemma A.1 (see Appendix A), $h_R < h_P$. Our preliminary remark yields that

$$h_P \leq h_P^* \qquad \text{and} \qquad h_R \geq h_R^*. \tag{22}$$

However, by strict concavity of $\pi(\cdot)$, $h_P^* > h_R^*$ for any $M > 0$.

As noted in the text, for the poor and the rich not to be tempted to deviate to the other group's activity, the following condition must hold:

$$\mu\Big(\frac{h_P}{h_0 + h_P} - \chi\Big)M \leq U(h_R) - U(h_P) + \pi(e^*(h_R) + M) - \pi(e^*(h_P)) \ \leq \ \mu\Big(\frac{h_R}{h_0 + h_R} - \chi\Big)M. \tag{23}$$

With our equilibrium concept, off-path deviations to an activity $h \in (h_P, h_R)$ are at-

[68] We use the continuity of all expressions with respect to $h \in (0, +\infty)$.

tributed to poor agents with probability 1 if for any belief $p \in \{0,1\}$,

$$U(h) + \pi(e^*(h) + pM) - \mu\Big(\frac{h}{h_0 + h} - \chi\Big)pM - U(h_P) - \pi(e^*(h_P))$$
$$> U(h) + \pi(e^*(h) + pM) + \frac{\mu h}{h_0 + h}M - \mu\Big(\frac{h}{h_0 + h} - \chi\Big)pM - U(h_R) - \pi(e^*(h) + M) - \mu\chi M,$$

i.e. if

$$\mu\Big(\frac{h}{h_0 + h} - \chi\Big)M < U(h_R) - U(h_P) + \pi(e^*(h_R) + M) - \pi(e^*(h_P)),$$

and to rich agents with probability 1 if

$$\mu\Big(\frac{h}{h_0 + h} - \chi\Big)M > U(h_R) - U(h_P) + \pi(e^*(h_R) + M) - \pi(e^*(h_P)),$$

Let h' be such that

$$\mu\Big(\frac{\mu h'}{h_0 + h'} - \chi\Big)M = U(h_R) - U(h_P) + \pi(e^*(h_R) + M) - \pi(e^*(h_P)).$$

Then, condition (23) implies that $h' \in [h_P, h_R]$, and the necessary and sufficient existence conditions for a separating equilibrium in pure strategies thus write as

$$\begin{cases} U(h_P) + \pi(e^*(h_P)) = \max\limits_{0 \le h \le h'} U(h) + \pi(e^*(h)), \\ U(h_R) + \pi(e^*(h_R)) = \max\limits_{h' \le h \le +\infty} U(h) + \pi(e^*(h) + M). \end{cases}$$

Therefore, two (mutually exclusive) cases arise:

(i) $h' > h_P^*$, and then $h_P = h_P^*$ and $h_R = h' > h_P^*$, i.e.

$$\mu\Big(\frac{h_R}{h_0 + h_R} - \chi\Big)M = U(h_R) - U(h_P^*) + \pi(e^*(h_R) + M) - \pi(e^*(h_P^*)).$$

(ii) $h' < h_R^*$, and then $h_R = h_R^*$ and $h_P = h' < h_R$, i.e.

$$\mu\Big(\frac{h_P}{h_0 + h_P} - \chi\Big)M = U(h_R^*) - U(h_P) + \pi(e^*(h_R^*) + M) - \pi(e^*(h_P)).$$

Indeed, if $h_R^* \le h' \le h_P^*$, then both the rich and the poor choose activity h', a contradiction.

Lastly, by strict concavity, $h_R > h_P^* > h_R^*$ implies that $U(h_R)+\pi(e^*(h_R)+M) < U(h_P^*)+\pi(e^*(h_P^*)+M)$, whereas $h_P < h_R^* < h_P^*$ implies that $U(h_P)+\pi(e^*(h_P) < U(h_R^*)+\pi(e^*(h_R^*))$.

Hence, case (i) corresponds to

$$\mu\Big(\frac{h_P^*}{h_0+h_P^*}-\chi\Big)M<\mu\Big(\frac{h_R}{h_0+h_R}-\chi\Big)M<\pi(e^*(h_P^*)+M)-\pi(e^*(h_P^*)),$$

("distinction"), while case (ii) corresponds to

$$\mu\Big(\frac{h_R^*}{h_0+h_R^*}-\chi\Big)M>\mu\Big(\frac{h_P}{h_0+h_P}-\chi\Big)M>\pi(e^*(h_R^*)+M)-\pi(e^*(h_R^*)),$$

("displacement").[69]

□

I Proofs of Lemma 5 and Proposition 6

Let us restrict our attention to pure strategies. Let us consider a candidate pooling equilibrium in which both the rich and the poor choose activity $h \in \mathbb{R}_+$ and exert effort $e^\dagger(M)$ and $e^\dagger(0)$ respectively.

For any performance $y \in \mathbb{R}$, let us denote by $\tilde{p}(h,y)$ the probability that the audience attributes to an agent being rich ($w = M$) upon observing a performance y in activity h. Hence, by applying Bayes' rule,

$$\begin{aligned}\tilde{p}(h,y) &\equiv \mathbb{E}[w=M|h,y]\\ &= \frac{pe^{-\frac{1}{2}\frac{h_0h}{h_0+h}(y-e^\dagger(M)-M)^2}}{pe^{-\frac{1}{2}\frac{h_0h}{h_0+h}(y-e^\dagger(M)-M)^2}+(1-p)e^{-\frac{1}{2}\frac{h_0h}{h_0+h}(y-e^\dagger(0))^2}}\\ &= \frac{pe^{-\frac{1}{2}\frac{h_0h}{h_0+h}[(M+e^\dagger(M))^2-e^\dagger(0)^2]}}{pe^{-\frac{1}{2}\frac{h_0h}{h_0+h}[(M+e^\dagger(M))^2-e^\dagger(0)^2]}+(1-p)e^{-\frac{h_0h}{h_0+h}y[M+e^\dagger(M)-e^\dagger(0)]}}.\end{aligned}$$

Given the other agents' efforts $e^\dagger(M)$, $e^\dagger(0)$, an agent's optimal effort (if any) in activity h maximizes

$$\begin{aligned}&\mathbb{E}_{\theta+\varepsilon_h}\left[\tilde{p}(h,\theta+e+w+\varepsilon_h)\right]\left[\Big(\mu\eta-\frac{\mu h}{h_0+h}\Big)[e^\dagger(M)-e^\dagger(0)]+\Big(\mu\chi-\frac{\mu h}{h_0+h}\Big)M\right]\\ &\quad+\Big(\mu\eta-\frac{\mu h}{h_0+h}\Big)e^\dagger(0)+\frac{\mu h}{h_0+h}(e+w)-g(e), \qquad (24)\end{aligned}$$

[69] Note that by strict concavity of $\pi(\cdot)$,

$$\pi(e^*(h_R^*)+M)-\pi(e^*(h_R^*))>\pi(e^*(h_P^*)+M)-\pi(e^*(h_P^*)),$$

and thus, as $h_P^*>h_R^*$,

$$\frac{\mu h_P^*}{h_0+h_R^*}M-\left[\pi(e^*(h_P^*)+M)-\pi(e^*(h_P^*))\right]>\frac{\mu h_R^*}{h_0+h_R^*}M-\left[\pi(e^*(h_R^*)+M)-\pi(e^*(h_R^*))\right].$$

where by construction, $\theta+\varepsilon_h$ is normally distributed with mean 0 and precision $h_0h/(h_0+h)$.

Characterization of the efforts $e^{\dagger}(M)$, $e^{\dagger}(0)$. In activity h, given the other agents' efforts $e^{\dagger}(w)$, as g is twice continuously differentiable with $g'(0)=0$ and $\lim_{e\to+\infty} g'(e)=+\infty$, an agent's optimal effort is either interior ($e\in(0,+\infty)$) or nil ($e=0$).

Let us derive the first-order conditions. To alleviate the notation, define $\varphi(e^{\dagger}(M),e^{\dagger}(0))$ as

$$\begin{aligned}\varphi(e^{\dagger}(M),e^{\dagger}(0)) \;&\equiv\; (1-p)\frac{h_0h}{h_0+h}\left[M+e^{\dagger}(M)-e^{\dagger}(0)\right]\\ &\qquad\times\left[\left(\mu\eta-\frac{\mu h}{h_0+h}\right)[e^{\dagger}(M)-e^{\dagger}(0)]+\left(\mu\chi-\frac{\mu h}{h_0+h}\right)M\right],\end{aligned}$$

and for $w\in\{0,M\}$ and $e\in\mathbb{R}_+$, $E(w,e)$ by

$$E(w,e) \;\equiv\; \mathbb{E}_{\theta+\varepsilon_h}\left[\frac{pe^{-\frac{1}{2}\frac{h_0h}{h_0+h}[(M+e^{\dagger}(M))^2-e^{\dagger}(0)^2]}e^{-\frac{h_0h}{h_0+h}(\theta+\varepsilon_h+e+w)[M+e^{\dagger}(M)-e^{\dagger}(0)]}}{\left(pe^{-\frac{1}{2}\frac{h_0h}{h_0+h}[(M+e^{\dagger}(M))^2-e^{\dagger}(0)^2]}+(1-p)e^{-\frac{h_0h}{h_0+h}(\theta+\varepsilon_h+e+w)[M+e^{\dagger}(M)-e^{\dagger}(0)]}\right)^2}\right].$$

Hence, given the audience's beliefs, if a rich agent's optimal effort $e(M)$ is interior, it satisfies the first-order condition

$$g'(e(M)) \;=\; \frac{\mu h}{h_0+h} \;+\; \varphi(e^{\dagger}(M),e^{\dagger}(0)).E(M,e(M)),$$

while similarly, if a poor agent's effort $e(0)$ is interior, it satisfies the first-order condition:

$$g'(e(0)) \;=\; \frac{\mu h}{h_0+h} \;+\; \varphi(e^{\dagger}(M),e^{\dagger}(0)).E(0,e(0)).$$

Suppose that the audience's beliefs are accurate and that the agents' optimal efforts are interior, and thus that

$$\begin{cases} g'(e^{\dagger}(M)) \;=\; \dfrac{\mu h}{h_0+h} \;+\; \varphi(e^{\dagger}(M),e^{\dagger}(0)).E(M,e^{\dagger}(M)),\\ g'(e^{\dagger}(0)) \;=\; \dfrac{\mu h}{h_0+h} \;+\; \varphi(e^{\dagger}(M),e^{\dagger}(0)).E(0,e^{\dagger}(0)).\end{cases} \tag{25}$$

Let $\Delta\equiv M+e^{\dagger}(M)-e^{\dagger}(0)$. Rearranging and using a change of variables yields that

$$E(M,e^{\dagger}(M)) \;=\; \frac{1}{\sqrt{2\pi}}\sqrt{\frac{h_0h}{h_0+h}}\int_{-\infty}^{\infty}\frac{pe^{-\frac{1}{2}\frac{h_0h}{h_0+h}(\theta+\Delta)^2}}{\left[p+(1-p)e^{-\frac{1}{2}\frac{h_0h}{h_0+h}(\Delta^2+2\theta\Delta)}\right]^2}d\theta,$$

and

$$E(0, e^{\dagger}(0)) = \frac{1}{\sqrt{2\pi}}\sqrt{\frac{h_0 h}{h_0+h}} \int_{-\infty}^{\infty} \frac{pe^{-\frac{1}{2}\frac{h_0 h}{h_0+h}(\theta+\Delta)^2}}{\left[pe^{-\frac{1}{2}\frac{h_0 h}{h_0+h}(\Delta^2+2\theta\Delta)} + (1-p)\right]^2} d\theta,$$

With a degenerate prior $p \in \{0, 1\}$, the above system implies that

$$e^{\dagger}(M) = e^{\dagger}(0) = \frac{\mu h}{h_0 + h} = e^*(h),$$

as emphasized earlier. By contrast, with a nondegenerate, interior prior $p \in (0, 1)$, whenever $e^{\dagger}(M)$ and $e^{\dagger}(0)$ are interior, they are either both strictly higher than $e^*(h)$ if $\varphi(e^{\dagger}(M), e^{\dagger}(0)) > 0$, and both strictly lower than $e^*(h)$ if $\varphi(e^{\dagger}(M), e^{\dagger}(0)) < 0$.

In addition, suppose p is either in a neighborhood of 0 or in a neighborhood of 1. Then, as the right-hand sides of (25) are strictly positive for p in such regions, the optimal efforts $e^{\dagger}(M)$ and $e^{\dagger}(0)$ are interior and given by (25). The implicit function theorem then yields that $e^{\dagger}(M)$ and $e^{\dagger}(0)$ are locally continuously differentiable with respect to p.[70] In particular, by differentiation in $p = 0$, using that $e^{\dagger}(M) = e^{\dagger}(0) = e^*(h)$ whenever $p \in \{0, 1\}$,

$$\begin{cases} g''(e^*(h)) \left.\dfrac{\partial e^{\dagger}(M)}{\partial p}\right|_{p=0} = \dfrac{h_0 h}{h_0+h}\left(\mu\chi - \dfrac{\mu h}{h_0+h}\right) M^2 \dfrac{1}{\sqrt{2\pi}}\sqrt{\dfrac{h_0 h}{h_0+h}} \displaystyle\int_{-\infty}^{\infty} \dfrac{e^{-\frac{1}{2}\frac{h_0 h}{h_0+h}(\theta+M)^2}}{e^{-\frac{h_0 h}{h_0+h}(M^2+2\theta M)}} d\theta \\ g''(e^*(h)) \left.\dfrac{\partial e^{\dagger}(0)}{\partial p}\right|_{p=0} = \dfrac{h_0 h}{h_0+h}\left(\mu\chi - \dfrac{\mu h}{h_0+h}\right) M^2 \dfrac{1}{\sqrt{2\pi}}\sqrt{\dfrac{h_0 h}{h_0+h}} \displaystyle\int_{-\infty}^{\infty} e^{-\frac{1}{2}\frac{h_0 h}{h_0+h}(\theta+M)^2} d\theta \end{cases}$$

and thus,

$$\begin{cases} g''(e^*(h)) \left.\dfrac{\partial e^{\dagger}(M)}{\partial p}\right|_{p=0} = \dfrac{h_0 h}{h_0+h}\left(\mu\chi - \dfrac{\mu h}{h_0+h}\right) M^2 e^{\frac{h_0 h}{h_0+h}M^2} \\ g''(e^*(h)) \left.\dfrac{\partial e^{\dagger}(0)}{\partial p}\right|_{p=0} = \dfrac{h_0 h}{h_0+h}\left(\mu\chi - \dfrac{\mu h}{h_0+h}\right) M^2, \end{cases} \tag{26}$$

[70] The determinant of the Jacobian matrix associated with (25) is equal to

$$\left[g''(e^{\dagger}(M)) - \frac{\partial}{\partial e^{\dagger}(M)}\Big(\varphi(e^{\dagger}(M), e^{\dagger}(0)).E(M, e^{\dagger}(M))\Big)\right]\left[g''(e^{\dagger}(0)) - \frac{\partial}{\partial e^{\dagger}(0)}\Big(\varphi(e^{\dagger}(M), e^{\dagger}(0)).E(0, e^{\dagger}(0))\Big)\right]$$
$$- \frac{\partial}{\partial e^{\dagger}(0)}\Big(\varphi(e^{\dagger}(M), e^{\dagger}(0)).E(M, e^{\dagger}(M))\Big)\frac{\partial}{\partial e^{\dagger}(M)}\Big(\varphi(e^{\dagger}(M), e^{\dagger}(0)).E(0, e^{\dagger}(0))\Big)$$

which is strictly positive for p in a neighborhood of 0 or 1.

and hence, for $h/(h_0+h) < \chi$,

$$\left.\frac{\partial e^{\dagger}(M)}{\partial p}\right|_{p=0} > \left.\frac{\partial e^{\dagger}(0)}{\partial p}\right|_{p=0} > 0,$$

while for $h/(h_0+h) > \chi$,

$$\left.\frac{\partial e^{\dagger}(M)}{\partial p}\right|_{p=0} < \left.\frac{\partial e^{\dagger}(0)}{\partial p}\right|_{p=0} < 0.$$

Similarly, differentiation in $p=1$ yields that

$$\begin{cases} g''(e^*(h))\left.\dfrac{\partial e^{\dagger}(M)}{\partial p}\right|_{p=1} = -\dfrac{h_0 h}{h_0+h}\left(\mu\chi - \dfrac{\mu h}{h_0+h}\right)M^2 \\ g''(e^*(h))\left.\dfrac{\partial e^{\dagger}(0)}{\partial p}\right|_{p=1} = -\dfrac{h_0 h}{h_0+h}\left(\mu\chi - \dfrac{\mu h}{h_0+h}\right)M^2 e^{\frac{h_0 h}{h_0+h}M^2}, \end{cases}$$

and hence, for $h/(h_0+h) < \chi$,

$$\left.\frac{\partial e^{\dagger}(0)}{\partial p}\right|_{p=1} < \left.\frac{\partial e^{\dagger}(M)}{\partial p}\right|_{p=1} < 0,$$

while for $h/(h_0+h) > \chi$,

$$\left.\frac{\partial e^{\dagger}(0)}{\partial p}\right|_{p=1} > \left.\frac{\partial e^{\dagger}(M)}{\partial p}\right|_{p=1} > 0.$$

This establishes the claim made in the text that the optimal effort level of a rich agent in activity h lies further away from $e^*(h)$ than the optimal effort level of a poor agent if $p(h)$ is in a neighborhood of 0, and closer to $e^*(h)$ if $p(h)$ is in a neighborhood of 1.

To show Lemma 5, we note that (25) imply that (a) $e^{\dagger}(M) = e^*(h)$ if and only if $e^{\dagger}(0) = e^*(h)$ (as each equality holds if and only if $\varphi(e^{\dagger}(M), e^{\dagger}(0)) = 0$), and that (b) there exists no $p \in (0,1)$ such that $e^{\dagger}(M) = e^{\dagger}(0) = e^*(h)$, which yields the result under the assumption that $e^{\dagger}(M)$, $e^{\dagger}(0)$ are continuous with respect to p. Indeed, $e^{\dagger}(M) = e^{\dagger}(0)$ implies that

$$\varphi(e^{\dagger}(M), e^{\dagger}(0)) = (1-p)\frac{h_0 h}{h_0+h}\left(\mu\chi - \frac{\mu h}{h_0+h}\right)M^2,$$

and thus, since $E(M, e^{\dagger}(M))$ and $E(0, e^{\dagger}(0))$ are strictly positive for $p \in (0,1)$, $e^{\dagger}(M)$ and $e^{\dagger}(0)$ are both strictly higher than $e^*(h)$ if $\chi > h/(h_0+h)$ (distinction) and both strictly lower than $e^*(h)$ if $\chi < h/(h_0+h)$ (displacement). (Lastly, as noted before, the continuity of $e^{\dagger}(M)$, $e^{\dagger}(0)$ with respect to p obtains for p in a neighborhood of 0 or 1 via the implicit function theorem.)

I.1 Proof of Proposition 6

Claim (i) follows immediately from Lemma D.2 in Appendix D.1. Let us show claim (ii).

(Candidate) Equilibrium payoffs For $w \in \{0, M\}$, let $p^\dagger(h,w)$ be the expectation formed by an agent with head start w of the probability with which the audience will consider that agent to be rich ($w = M$) after observing their choice of activity and performance:

$$p^\dagger(h,w) = \mathbb{E}_{\theta+\varepsilon_h}\left[\tilde{p}(h;\theta + e^\dagger(w) + w + \varepsilon_h)\right].$$

Hence,

$$p^\dagger(h,M) = \mathbb{E}_{\theta+\varepsilon_h}\left[\frac{pe^{-\frac{1}{2}\frac{h_0h}{h_0+h}[(M+e^\dagger(M))^2-e^\dagger(0)^2]}}{pe^{-\frac{1}{2}\frac{h_0h}{h_0+h}[(M+e^\dagger(M))^2-e^\dagger(0)^2]}+(1-p)e^{-\frac{h_0h}{h_0+h}(\theta+\varepsilon_h+M+e^\dagger(M))[M+e^\dagger(M)-e^\dagger(0)]}}\right]$$

$$p^\dagger(h,0) = \mathbb{E}_{\theta+\varepsilon_h}\left[\frac{pe^{-\frac{1}{2}\frac{h_0h}{h_0+h}[(M+e^\dagger(M))^2-e^\dagger(0)^2]}}{pe^{-\frac{1}{2}\frac{h_0h}{h_0+h}[(M+e^\dagger(M))^2-e^\dagger(0)^2]}+(1-p)e^{-\frac{h_0h}{h_0+h}(\theta+\varepsilon_h+e^\dagger(0))[M+e^\dagger(M)-e^\dagger(0)]}}\right]$$

Let $\Delta \equiv M + e^\dagger(M) - e^\dagger(0)$. By construction,

$$p^\dagger(h,M) = \frac{1}{\sqrt{2\pi}}\sqrt{\frac{h_0h}{h_0+h}}\int_{-\infty}^{\infty}\frac{pe^{-\frac{1}{2}\frac{h_0h}{h_0+h}\theta^2}}{p+(1-p)e^{-\frac{1}{2}\frac{h_0h}{h_0+h}(2\theta\Delta+\Delta^2)}}d\theta$$

$$p^\dagger(h,0) = \frac{1}{\sqrt{2\pi}}\sqrt{\frac{h_0h}{h_0+h}}\int_{-\infty}^{\infty}\frac{pe^{-\frac{1}{2}\frac{h_0h}{h_0+h}\theta^2}}{p+(1-p)e^{-\frac{1}{2}\frac{h_0h}{h_0+h}(2\theta\Delta-\Delta^2)}}d\theta,$$

and thus by a change of variables,

$$p^\dagger(h,M) = \frac{1}{\sqrt{2\pi}}\sqrt{\frac{h_0h}{h_0+h}}\int_{-\infty}^{\infty}\frac{pe^{-\frac{1}{2}\frac{h_0h}{h_0+h}\left(\theta-\frac{\Delta}{2}\right)^2}}{p+(1-p)e^{-\frac{h_0h}{h_0+h}\theta\Delta}}d\theta \tag{27}$$

$$p^\dagger(h,0) = \frac{1}{\sqrt{2\pi}}\sqrt{\frac{h_0h}{h_0+h}}\int_{-\infty}^{\infty}\frac{pe^{-\frac{1}{2}\frac{h_0h}{h_0+h}\left(\theta+\frac{\Delta}{2}\right)^2}}{p+(1-p)e^{-\frac{h_0h}{h_0+h}\theta\Delta}}d\theta. \tag{28}$$

The expected payoff in activity h of an agent with head start $w \in \{0, M\}$, denoted by $v(w)$, is thus given by

$$v(w) = p^\dagger(h,w)\left[\left(\mu\eta - \frac{\mu h}{h_0+h}\right)[e^\dagger(M)-e^\dagger(0)] + \left(\mu\chi - \frac{\mu h}{h_0+h}\right)M\right] + \left(\mu\eta - \frac{\mu h}{h_0+h}\right)e^\dagger(0)$$
$$+\frac{\mu h}{h_0+h}(e^\dagger(w)+w) - g(e^\dagger(w)).$$

Hence, the expected payoffs of rich and poor agents in activity h are given respectively by

$$v(M) = -(1-p^{\dagger}(h,M))\Big[\Big(\mu\eta - \frac{\mu h}{h_0+h}\Big)[M+e^{\dagger}(M)-e^{\dagger}(0)] + (\mu\chi - \mu\eta)M\Big] + \mu\chi M \\ + \mu\eta e^{\dagger}(M) - g(e^{\dagger}(M))$$

for a rich agent, and by

$$v(0) = p^{\dagger}(h,0)\Big[\Big(\mu\eta - \frac{\mu h}{h_0+h}\Big)[M+e^{\dagger}(M)-e^{\dagger}(0)] + (\mu\chi - \mu\eta)M\Big] + \mu\eta e^{\dagger}(0) - g(e^{\dagger}(0))$$

for a poor agent. To alleviate the notation, let

$$\begin{aligned}\phi(h,\eta,\chi) &\equiv \Big(\mu\eta - \frac{\mu h}{h_0+h}\Big)[M+e^{\dagger}(M)-e^{\dagger}(0)] + (\mu\chi - \mu\eta)M \\ &= \Big(\mu\eta - \frac{\mu h}{h_0+h}\Big)\Delta + (\mu\chi - \mu\eta)M,\end{aligned}$$

so that the expected payoffs of rich and poor agents in activity h write as

$$\begin{cases} v(M) = -(1-p^{\dagger}(h,M))\phi(h,\eta,\chi) + \mu\chi M + \mu\eta e^{\dagger}(M) - g(e^{\dagger}(M)) \\ v(0) = p^{\dagger}(h,0)\phi(h,\eta,\chi) + \mu\eta e^{\dagger}(0) - g(e^{\dagger}(0)). \end{cases} \tag{29}$$

With degenerate off-path beliefs, the optimal effort of both rich and poor agents in an off-path activity h' is equal to $e^*(h')$ (see the proof of Lemma 5, Appendix I). Hence, the expected payoff of an agent with headstart w when choosing the off-path activity h' is equal to

$$p(h')\Big(\mu\chi - \frac{\mu h'}{h_0+h'}\Big)M + \frac{\mu h'}{h_0+h'}w + \mu\eta e^*(h') - g(e^*(h'))$$

with $p(h') \in \{0,1\}$ the (degenerate) off-path belief that an agent choosing activity h' be rich ($w = M$).

Consequently, with our equilibrium concept (see Appendix A for the formal definition), a deviation to h' is attributed to a rich agent with probability 1 if

$$\frac{\mu h'}{h_0+h'}M - v(M) + v(0) > 0, \tag{30}$$

and the deviation is attributed to a poor agent with probability 1 if the opposite inequality holds strictly.

Proof of Claim (ii). Let p be in a neighborhood of 0. Consider a candidate pooling equilibrium in pure strategies in activity h. By the previous arguments, the optimal efforts $e^\dagger(M)$, $e^\dagger(0)$ within activity h are well-defined and interior, given by the first-order condition (25). Moreover, as p is in a neighborhood of 0, (26) implies that

$$\begin{cases} e^\dagger(M) = e^*(h) + \dfrac{1}{g''(e^*(h))}\dfrac{h_0 h}{h_0+h}\left(\mu\chi - \dfrac{\mu h}{h_0+h}\right)M^2 e^{\frac{h_0 h}{h_0+h}M^2} p + O(p^2), \\ e^\dagger(0) = e^*(h) + \dfrac{1}{g''(e^*(h))}\dfrac{h_0 h}{h_0+h}\left(\mu\chi - \dfrac{\mu h}{h_0+h}\right)M^2 p + O(p^2), \end{cases}$$

and thus

$$e^\dagger(M) - e^\dagger(0) = \frac{1}{g''(e^*(h))}\frac{h_0 h}{h_0+h}\left(\mu\chi - \frac{\mu h}{h_0+h}\right)M^2\left[e^{\frac{h_0 h}{h_0+h}M^2} - 1\right]p + O(p^2). \tag{31}$$

Hence,[71]

$$\begin{aligned} \phi(h,\eta,\chi) =& \left(\mu\chi - \frac{\mu h}{h_0+h}\right)M \\ &+ \left(\mu\eta - \frac{\mu h}{h_0+h}\right)\frac{1}{g''(e^*(h))}\frac{h_0 h}{h_0+h}\left(\mu\chi - \frac{\mu h}{h_0+h}\right)M^2\left[e^{\frac{h_0 h}{h_0+h}M^2} - 1\right]p + O(p^2). \end{aligned}$$

Furthermore, with the above arguments, for p in a neighborhood of 0, by differentiating (27) in $p = 0$,

$$\begin{cases} p^\dagger(h,M) = \dfrac{p}{\sqrt{2\pi}}\sqrt{\dfrac{h_0 h}{h_0+h}}\displaystyle\int_{-\infty}^{\infty} e^{-\frac{1}{2}\frac{h_0 h}{h_0+h}\left(\theta - \frac{M}{2}\right)^2} e^{\frac{h_0 h}{h_0+h}\theta M} d\theta + O(p), \\ p^\dagger(h,0) = \dfrac{p}{\sqrt{2\pi}}\sqrt{\dfrac{h_0 h}{h_0+h}}\displaystyle\int_{-\infty}^{\infty} e^{-\frac{1}{2}\frac{h_0 h}{h_0+h}\left(\theta + \frac{M}{2}\right)^2} e^{\frac{h_0 h}{h_0+h}\theta M} d\theta + O(p) \end{cases}$$

and thus

$$\begin{cases} p^\dagger(h,M) = e^{\frac{h_0 h}{h_0+h}M^2} p + O(p), \\ p^\dagger(h,0) = p + O(p). \end{cases}$$

[71]By (31), for p in a neighborhood of 0,

$$\Delta = M + \frac{1}{g''(e^*(h))}\frac{h_0 h}{h_0+h}\left(\mu\chi - \frac{\mu h}{h_0+h}\right)M^2\left[e^{\frac{h_0 h}{h_0+h}M^2} - 1\right]p + O(p^2).$$

Consequently, writing (29) for p in a neighborhood of 0, the expected payoff of rich and poor agents in activity h are equal to

$$\begin{aligned}
v(M) = & \frac{\mu h}{h_0+h}M + \mu\eta e^*(h) - g(e^*(h)) + e^{\frac{h_0 h}{h_0+h}M^2}\left(\mu\chi - \frac{\mu h}{h_0+h}\right)Mp \\
& - \left(\mu\eta - \frac{\mu h}{h_0+h}\right)\frac{1}{g''(e^*(h))}\frac{h_0 h}{h_0+h}\left(\mu\chi - \frac{\mu h}{h_0+h}\right)M^2\left[e^{\frac{h_0 h}{h_0+h}M^2} - 1\right]p \\
& + \frac{\mu\eta - g'(e^*(h))}{g''(e^*(h))}\frac{h_0 h}{h_0+h}\left(\mu\chi - \frac{\mu h}{h_0+h}\right)M^2 e^{\frac{h_0 h}{h_0+h}M^2}p + O(p^2) \\
= & \frac{\mu h}{h_0+h}M + \mu\eta e^*(h) - g(e^*(h)) \\
& + e^{\frac{h_0 h}{h_0+h}M^2}\left(\mu\chi - \frac{\mu h}{h_0+h}\right)\left[1 + \left(\mu\eta - \frac{\mu h}{h_0+h}\right)\frac{1}{g''(e^*(h))}\frac{h_0 h}{h_0+h}M\right]Mp + O(p^2)
\end{aligned}$$

and

$$\begin{aligned}
v(0) = & \mu\eta e^*(h) - g(e^*(h)) + \left(\mu\chi - \frac{\mu h}{h_0+h}\right)Mp \\
& + \frac{\mu\eta - g'(e^*(h))}{g''(e^*(h))}\frac{h_0 h}{h_0+h}\left(\mu\chi - \frac{\mu h}{h_0+h}\right)M^2 p + O(p^2) \\
= & \mu\eta e^*(h) - g(e^*(h)) \\
& + \left(\mu\chi - \frac{\mu h}{h_0+h}\right)\left[1 + \left(\mu\eta - \frac{\mu h}{h_0+h}\right)\frac{1}{g''(e^*(h))}\frac{h_0 h}{h_0+h}M\right]Mp + 0(p^2).
\end{aligned}$$

Therefore, for p in a neighborhood of 0,

$$\begin{aligned}
& v(M) - v(0) \\
& = \frac{\mu h}{h_0+h}M + \left(e^{\frac{h_0 h}{h_0+h}M^2} - 1\right)\left(\mu\chi - \frac{\mu h}{h_0+h}\right)\left[1 + \left(\mu\eta - \frac{\mu h}{h_0+h}\right)\frac{1}{g''(e^*(h))}\frac{h_0 h}{h_0+h}M\right]Mp + O(p^2),
\end{aligned}$$

and thus, by (30), a deviation to activity $h' \neq h$ is attributed to a rich agent with probability 1 if $h' > h$, and to a poor agent with probability 1 if $h' < h$.

Let us investigate deviations. Suppose that $h/(h_0+h) < \eta$ and consider the deviation of rich agents to activity h' such that $h' > h$. A rich agent's differential payoff from the deviation is equal to

$$\begin{aligned}
& \left(\mu\chi - \frac{\mu h}{h_0+h}\right)M + \mu\eta e^*(h') - g(e^*(h')) - \mu\eta e^*(h) + g(e^*(h)) \\
& \quad - e^{\frac{h_0 h}{h_0+h}M^2}\left(\mu\chi - \frac{\mu h}{h_0+h}\right)\left[1 + \left(\mu\eta - \frac{\mu h}{h_0+h}\right)\frac{1}{g''(e^*(h))}\frac{h_0 h}{h_0+h}M\right]Mp + O(p^2).
\end{aligned}$$

Hence, a necessary equilibrium condition is that $\chi < h/(h_0+h)$ (as otherwise, by strict convexity of g, a deviation to h^* such that $h^*/(h_0+h^*) = \eta$ would be strictly profitble for

p in a neighborhood of 0). Similarly, a rich agent's differential payoff from a deviation to an activity $h' < h$ is equal to

$$\left(\frac{\mu h'}{h_0 + h'} - \frac{\mu h}{h_0 + h}\right)M + \mu\eta e^*(h') - g(e^*(h')) - \mu\eta e^*(h) + g(e^*(h))$$
$$- e^{\frac{h_0 h}{h_0+h}M^2}\left(\mu\chi - \frac{\mu h}{h_0 + h}\right)\left[1 + \left(\mu\eta - \frac{\mu h}{h_0 + h}\right)\frac{1}{g''(e^*(h))}\frac{h_0 h}{h_0 + h}M\right]Mp + O(p^2).$$

In particular, if $h'/(h_0 + h') = h/(h_0 + h) + O(p^2)$, the differential payoff from the deviation is equal to

$$-e^{\frac{h_0 h}{h_0+h}M^2}\left(\mu\chi - \frac{\mu h}{h_0 + h}\right)\left[1 + \left(\mu\eta - \frac{\mu h}{h_0 + h}\right)\frac{1}{g''(e^*(h))}\frac{h_0 h}{h_0 + h}M\right]Mp + O(p^2),$$

which is strictly positive as $\chi < h/(h_0 + h) < \eta$.

Suppose alternatively that $h/(h_0 + h) > \eta$. A poor agent's differential payoff from the deviation is equal to

$$\mu\eta e^*(h') - g(e^*(h')) - \mu\eta e^*(h) + g(e^*(h))$$
$$- \left(\mu\chi - \frac{\mu h}{h_0 + h}\right)\left[1 + \left(\mu\eta - \frac{\mu h}{h_0 + h}\right)\frac{1}{g''(e^*(h))}\frac{h_0 h}{h_0 + h}M\right]Mp + O(p^2),$$

which is strictly positive for $h'/(h_0 + h') = \eta$ and p in a neighborhood of 0.

Lastly, suppose that $h/(h_0 + h) = \eta$. A rich agent's deviation to $h' < h$ such that $h'/(h_0 + h') = h/(h_0 + h) + O(p^2)$ yields a differential payoff equal to

$$-e^{\frac{h_0 h}{h_0+h}M^2}\left(\mu\chi - \frac{\mu h}{h_0 + h}\right)Mp + O(p^2),$$

and therefore, a necessary equilibrium condition is that $\chi \geq h/(h_0 + h) = \eta$. However, a rich agent's deviation to $h' > h$ such that $h'/(h_0 + h') = h/(h_0 + h) + O(p^2)$ yields a differential payoff equal to

$$\left(\mu\chi - \frac{\mu h}{h_0 + h}\right)M - e^{\frac{h_0 h}{h_0+h}M^2}\left(\mu\chi - \frac{\mu h}{h_0 + h}\right)Mp + O(p^2),$$

which implies that a necessary equilibrium condition is that $\chi \leq h/(h_0 + h) = \eta$.

Hence, for p in a neighborhood of 0, a necessary equilibrium condition is that $\chi = \eta$. The same necessary equilibrium condition obtains similarly for p in a neighborhood of 1. Claim (ii) follows.

J Proof of Proposition 7

Suppose agents choose among a continuum of activities indexed by their precision $h \in \mathbb{R}_+$, with period-t transfer β_t constant across activities, and a quadratic cost of effort $g(e) = e^2/2$. We denote by $\hat{e}_t(I)$ the audience's expectation of the effort level $e_t \in \mathbb{R}_+$ that the agent exerts in period $t \in \{1, 2\}$, given the set of public observables $I \subset \{h_1, y_1, h_2, y_2\}$. Similarly, we denote by $\hat{w}(I)$ the audience's expectation of the agent's head start w conditional on observables I. We look for equilibria in pure strategies, with degenerate off-path beliefs for the audience.

No head-start inequality. Suppose first that there is no head-start inequality. We assume that the audience has *passive beliefs* regarding the agent's period-1 effort, i.e. only uses h_1 to form its belief about e_1, and in particular, does not update its belief after observing y_1, h_2 and y_2.

Consequently, for any h_1, y_1, e_1 and audience's on-path belief $\hat{e}_1(h_1)$, the agent's period-2 activity choice $h_2^\dagger(h_1, y_1, e_1, \hat{e}_1(h_1))$ is a solution, if any, to

$$\max_{h_2} \left(\frac{\mu h_1}{h_0 + h_1 + h_2} [y_1 - \hat{e}_1(h_1)] + \mu\eta\hat{e}_1(h_1) \right. \\ \left. + \frac{\mu h_2}{h_0 + h_1 + h_2} \frac{h_1}{h_0 + h_1} [y_1 - e_1] + \mu\eta e_2^*(h_1, h_2) - g(e_2^*(h_1, h_2)) \right)$$

where $e_2^*(h_1, h_2)$ is given by

$$g'(e_2^*) = \frac{\mu h_2}{h_0 + h_1 + h_2}.$$

Hence, the agent's period-2 activity choice is a solution, if any, to[72]

$$\max_{h_2} \left(\frac{\mu h_1}{h_0 + h_1 + h_2} [e_1 - \hat{e}_1(h_1)] + \mu\eta e_2^*(h_1, h_2) - g(e_2^*(h_1, h_2)) \right) \tag{32}$$

A solution $h_2^\dagger < +\infty$ exists if and only if $\mu\eta + [h_1/(h_0 + h_1)][\hat{e}_1(h_1) - e_1] < \mu$. Then, the objective being continuously differentiable, first strictly increasing then strictly decreasing

[72] We use in particular that

$$\frac{\mu h_2}{h_0 + h_1 + h_2} \frac{h_1}{h_0 + h_1} = \left(1 - \frac{\mu(h_0 + h_1)}{h_0 + h_1 + h_2}\right) \frac{h_1}{h_0 + h_1}.$$

with respect to h_2, whenever interior, $h_2^\dagger$ is uniquely given by the first-order condition:

$$\frac{\mu h_2^\dagger}{h_0+h_1+h_2^\dagger} = \mu\eta + \frac{h_1}{h_0+h_1}[\hat{e}_1(h_1)-e_1],$$

and $h_2^\dagger = 0$ whenever $\mu\eta + [h_1/(h_0+h_1)][\hat{e}_1(h_1)-e_1] < 0$. In particular, as long as it remains interior, $h_2^\dagger$ strictly decreases with e_1.

Given a continuation strategy $(h_2^\dagger(h_1, e_1))_{e_1\geq 0}$, the agent's period-1 effort $e_1^\dagger(e_1, \hat{e}_1(h_1))$ is then a solution, if any, to

$$\max_{e_1} \mathbb{E}_{\theta+\varepsilon_1}\Bigg[\frac{\mu h_1}{h_0+h_1+h_2^\dagger}[\theta+\varepsilon_1+e_1-\hat{e}_1(h_1)] + \mu\eta\hat{e}_1(h_1) - \delta^{-1}g(e_1) + \frac{\mu h_2^\dagger}{h_0+h_1+h_2^\dagger}\frac{h_1}{h_0+h_1}[\theta+\varepsilon_1] + \mu\eta e_2^*(h_1,h_2^\dagger) - g(e_2^*(h_1,h_2^\dagger))\Bigg]$$

i.e. to

$$= \max_{e_1}\left(\frac{\mu h_1}{h_0+h_1+h_2^\dagger}[e_1-\hat{e}_1(h_1)] - \delta^{-1}g(e_1) + \mu\eta e_2^*(h_1,h_2^\dagger) - g(e_2^*(h_1,h_2^\dagger))\right) \tag{33}$$

as neither $\hat{e}_1(h_1)$ nor $h_2^\dagger$ depend on the realization of $\theta+\varepsilon_1$.[73] As $h_2^\dagger$ is a solution to (32), the agent's period-1 effort $e_1^\dagger$ is thus (uniquely) given by

$$e_1^\dagger(h_1, \hat{e}(h_1)) = \frac{\delta\mu h_1}{h_0+h_1+h_2^\dagger(h_1,e_1,\hat{e}(h_1))}$$

A necessary equilibrium condition is that $e_1^\dagger(h_1,\hat{e}(h_1)) = \hat{e}_1(h_1)$. Hence, under our assumptions, in equilibrium (if any), the period-2 activity choice h_2^* is (uniquely) given by

$$\frac{h_2^*}{h_0+h_1+h_2^*} = \eta,$$

and thus does not depend on e_1 nor on $\hat{e}_1$. In addition, h_2^* is strictly increasing and continuously differentiable with respect to h_1. As a consequence, in equilibrium (if any), the

[73]For $e_1 > 0$, the first derivative of the objective with respect to e_1 is equal to

$$\begin{aligned}&\frac{\mu h_1}{h_0+h_1+h_2^\dagger} - \delta^{-1}g'(e_1) + \left(\frac{\mu h_1}{(h_0+h_1+h_2^\dagger)^2}\big[\hat{e}_1(h_1)-e_1\big] + \big[\mu\eta - g'(e_2^*(h_1,h_2^\dagger))\big]\frac{\partial e_2^*}{\partial h_2}\right)\frac{\partial h_2^\dagger}{\partial e_1}\\ =&\frac{\mu h_1}{h_0+h_1+h_2^\dagger} - \delta^{-1}g'(e_1),\end{aligned}$$

as $h_2^\dagger$ is a solution to (32), while the second derivative of the objective with respect to e_1 is thus equal to

$$-\frac{\mu h_1}{(h_0+h_1+h_2^\dagger)^2}\frac{\partial h_2^\dagger}{\partial e_1} - \delta^{-1}g''(e_1) = \left(\frac{h_1}{h_0+h_1}\right)^2 - \delta^{-1} < 0.$$

period-1 effort choice e_1^* is (uniquely) given by

$$e_1^* = \frac{\delta\mu h_1}{h_0 + h_1 + h_2^*(h_1)}$$

Hence, let us check that these beliefs and strategies form an equilibrium of the continuation game starting after period-1 activity choice (h_1). Let us take the audience's belief on the agent's period-1 effort after observing h_1, as the degenerate belief putting probability 1 on $e_1^*(h_1)$. Hence,

$$\hat{e}_1(h_1) = \frac{\delta\mu h_1}{h_0 + h_1 + h_2^*(h_1)} = \frac{\delta\mu h_1}{(h_0 + h_1) + (h_0 + h_1)\eta/(1-\eta)} = \frac{\delta\mu(1-\eta)h_1}{h_0 + h_1}.$$

Then, the agent's objective when choosing h_2 strictly increases, resp. strictly decreases, for any $h_2 < h_2'$, resp. $h_2 > h_2'$ where

$$\begin{aligned}\frac{\mu h_2'}{h_0 + h_1 + h_2'} &= \mu\eta + \frac{h_1}{h_0 + h_1}\left(\frac{\delta\mu(1-\eta)h_1}{h_0 + h_1} - e_1\right) \\ &< \mu\eta + \delta\mu(1-\eta)\left(\frac{h_1}{h_0 + h_1}\right)^2 < \mu,\end{aligned}$$

and thus $h_2' < \infty$, i.e. there exists a (finite) solution to (32) and $h_2^\dagger$ is interior or nil when the audience's belief is given by $\hat{e}_1(h_1)$. By the above computations, whenever interior, $h_2^\dagger$ satisfies:

$$h_2^\dagger\left(\mu(1-\eta) - \frac{h_1}{h_0 + h_1}\left[\hat{e}_1(h_1) - e_1\right]\right) = (h_0 + h_1)\left(\mu\eta + \frac{h_1}{h_0 + h_1}\left[\hat{e}_1(h_1) - e_1\right]\right)$$

And thus

$$\begin{aligned}e_1^\dagger &= \frac{\delta\mu h_1}{h_0 + h_1 + h_2^\dagger(h_1, e_1^\dagger, \hat{e}(h_1))} \\ &= \frac{\delta\mu h_1}{h_0 + h_1}\frac{\mu(1-\eta) - \frac{h_1}{h_0 + h_1}\left[\hat{e}_1(h_1) - e_1\right]}{\mu(1-\eta) - \frac{h_1}{h_0 + h_1}\left[\hat{e}_1(h_1) - e_1\right] + \mu\eta + \frac{h_1}{h_0 + h_1}\left[\hat{e}_1(h_1) - e_1\right]} \\ &= \frac{\delta\mu h_1}{h_0 + h_1}\frac{\mu(1-\eta) - \frac{h_1}{h_0 + h_1}\left[\hat{e}_1(h_1) - e_1^\dagger\right]}{\mu}\end{aligned}$$

i.e. rearranging and replacing $\hat{e}_1(h_1)$ by its explicit expression,

$$\left[1 - \delta\left(\frac{h_1}{h_0 + h_1}\right)^2\right]e_1^\dagger = \delta\mu(1-\eta)\frac{\delta\mu h_1}{h_0 + h_1}\left[1 - \delta\left(\frac{h_1}{h_0 + h_1}\right)^2\right],$$

and thus $e_1^\dagger(h_1, \hat{e}_1(h_1)) = \hat{e}_1(h_1)$.

Similarly, $h_2^\dagger$ is not interior, thus $h_2^\dagger = 0$, if and only if

$$\mu\eta + \frac{h_1}{h_0 + h_1}[\hat{e}_1(h_1) - e_1] < 0.$$

But then,

$$e_1^\dagger = \frac{\delta\mu h_1}{h_0 + h_1},$$

and thus

$$\mu\eta + \frac{h_1}{h_0 + h_1}[\hat{e}_1(h_1) - e_1^\dagger] = \mu\eta\left(1 - \frac{\delta h_1^2}{(h_0 + h_1)^2}\right) > 0,$$

a contradiction. Therefore, the above strategies and beliefs form an equilibrium of the continuation game starting after period-1 activity choice: after choosing h_1, an agent chooses $e_1^*(h_1)$ and then $h_2^*(h_1)$, and the audience has (degenerate) belief $\hat{e}_1(h_1) = e_1^*(h_1)$ about the agent's period-1 effort. Under our assumptions, it is the unique equilibrium in pure strategies of the continuation game.

At the beginning of period 1, the agent thus chooses their activity h_1 by solving

$$\max_{h_1} \Big(\mu\eta e_1^*(h_1) - \delta^{-1} g(e_1^*(h_1)) + \mu\eta e_2^*(h_1, h_2^*(h_1)) - g(e_2^*(h_1, h_2^*(h_1)))\Big).$$

By construction, for any h_1,

$$e_2^*(h_1, h_2^*(h_1)) = (g')^{-1}(\mu\eta),$$

and thus $\mu\eta e_2^*(h_1, h_2^*(h_1)) - g(e_2^*(h_1, h_2^*(h_1)))$ does not depend on h_1. Hence, the agent's objective when choosing h_1 being strictly concave and continuously differentiable with respect to h_1, in equilibrium the agent chooses h_1^* such that

$$g'(e_1^*) = \eta, \qquad \text{i.e.} \qquad \frac{h_1^*}{h_0 + h_1^* + h_2^*(h_1^*)} = \eta,$$

which yields a unique solution h_1^*.

Head-start inequality. Let us now introduce head-start inequality. We begin by studying the agents' optimal activity and effort choices after any given history.

We look for fully separating equilibria in pure strategies. As a consequence, in any such equilibrium, the audience's expectation of the agent's head starts does not depend on realized performances y_1, y_2. We denote by $\hat{w}(h_1, h_2)$ the audience's expectation of the agent's head start after observing h_1, h_2.

Period-2 effort. Therefore, by linearity, for a choice of period-1 activity h_1, period-1 effort e_1 and period-2 activity h_2, an agent's optimal effort in period-2 solves[74]

$$\max_e \frac{\mu h_2}{h_0 + h_1 + h_2} e - g(e),$$

regardless of the agent's headstart, and is thus given by $e_2^*(h_1, h_2)$ as in the absence of headstart inequality.

Beliefs. Let us focus on monotone, degenerate beliefs, i.e. such that $\hat{w}(h_1, h_2) \in \{0, M\}$ and (weakly) increases in both its arguments. Hence, for any h_1, there exists a cutoff $h_2^c(h_1) \in [0, +\infty]$ such that $\hat{w}(h_1, h_2) = 0$ for any $h_2 < h_2^c$, and $\hat{w}(h_1, h_2)M$ for any $h_2 > h_2^c$. More specifically, we will later restrict our attention to (monotone and degenerate) beliefs $\hat{w}$ such that for any $h_1 \geq 0$, $h_2^c(h_1) \equiv h_2^*(h_1)$. In words, any choice of period-2 precision strictly above $h_2^*(h_1)$ is attributed to a rich agent ($\hat{w}(h_1, h_2) = M$), while any choice below is attributed to a poor agent ($\hat{w}(h_1, h_2) = 0$).

Choice of period-2 activity. After a period-1 activity choice h_1 and period-1 effort choice

[74]Indeed, an agent's optimal effort in period-2 solves

$$\max_{e_2} \Bigg(\frac{\mu h_1}{h_0 + h_1 + h_2} \big[y_1 - \hat{e}_1(h_1) - \hat{w}(h_1, h_2) \big] + \mu \eta \hat{e}_1(h_1) + \mu \chi \hat{w}(h_1, h_2)$$
$$+ \frac{\mu h_2}{h_0 + h_1 + h_2} \left(\frac{h_1}{h_0 + h_1} \big[y_1 - e_1 - w \big] + e_2 + w - e_2^*(h_1, h_2) - \hat{w}(h_1, h_2) \right) + \mu \eta e_2^*(h_1, h_2) - g(e_2) \Bigg).$$

e_1, the agent chooses h_2 by solving[75]

$$\begin{aligned}
&\max_{h_2}\Bigg(\frac{\mu h_1}{h_0+h_1+h_2}[y_1-\hat{e}_1(h_1)-\hat{w}(h_1,h_2)]+\mu\eta\hat{e}_1(h_1)+\mu\chi\hat{w}(h_1,h_2)\\
&\qquad+\frac{\mu h_2}{h_0+h_1+h_2}\left(\frac{h_1}{h_0+h_1}[y_1-e_1-w]+w-\hat{w}(h_1,h_2)\right)\\
&\qquad\qquad+\mu\eta e_2^*(h_1,h_2)-g(e_2^*(h_1,h_2))\Bigg)\\
&=\max_{h_2}\Bigg(\frac{\mu h_1}{h_0+h_1+h_2}[e_1-\hat{e}_1(h_1)]+\frac{\mu(h_1+h_2)}{h_0+h_1+h_2}[w-\hat{w}_1(h,h_2)]\\
&\qquad\qquad+\mu\eta e_2^*(h_1,h_2)-g(e_2^*(h_1,h_2))\Bigg)\\
&\equiv\max_{h_2}\Psi(h_1,e_1,h_2).
\end{aligned}$$

For simplicity, to ensure the existence of a maximum, we assume that there exists a perfectly revealing activity with infinite precision $h_2=+\infty$.

Let us distinguish two cases: $h_2\in(0,h_2^*(h_1))$ and $h_2\in(h_2^*(h_1),+\infty)$. On each of these open sets, the objective Ψ being continuously differentiable with respect to h_2, first strictly increasing then strictly decreasing. Whenever interior, the precisions that maximize the objective on each of these sets, denoted by $h_2^-(e_1,h_1)$ for the set $(0,h_2^*(h_1))$ and by $h_2^+(e_1,h_1)$ for the set $(h_2^*(h_1),+\infty)$, are uniquely given by the first-order condition:[76]

$$\begin{aligned}
\frac{\mu h_2^-}{h_0+h_1+h_2^-}&=\mu\eta+\frac{h_0}{h_0+h_1}w+\frac{h_1}{h_0+h_1}(\hat{e}_1(h_1)-e_1),\\
\frac{\mu h_2^+}{h_0+h_1+h_2^+}&=\mu\eta+\frac{h_0}{h_0+h_1}(w-M)+\frac{h_1}{h_0+h_1}(\hat{e}_1(h_1)-e_1),
\end{aligned}\tag{34}$$

In addition, let

$$h_2^-(e_1,h_1)\equiv\begin{cases}0 & \text{if } \mu\eta+\frac{h_0}{h_0+h_1}w+\frac{h_1}{h_0+h_1}(\hat{e}_1(h_1)-e_1)<0,\\ h_2^*(h_1) & \text{if } \frac{h_0}{h_0+h_1}w+\frac{h_1}{h_0+h_1}(\hat{e}_1(h_1)-e_1)\geq 0.\end{cases}\tag{35}$$

[75]The equality follows in particular from:

$$\frac{\mu h_2}{h_0+h_1+h_2}\frac{h_1}{h_0+h_1}=\left(1-\frac{\mu(h_0+h_1)}{h_0+h_1+h_2}\right)\frac{h_1}{h_0+h_1}.$$

[76]Recall that, by definition of $h_2^*(h_1)$,

$$\frac{\mu h_2^*(h_1)}{h_0+h_1+h_2^*(h_1)}=\mu\eta.$$

Similarly, let

$$h_2^+(e_1,h_1) \equiv \begin{cases} h_2^*(h_1) & \text{if } \dfrac{h_0}{h_0+h_1}(w-M) + \dfrac{h_1}{h_0+h_1}(\hat{e}_1(h_1)-e_1) \leq 0, \\ +\infty & \text{if } \mu\eta + \dfrac{h_0}{h_0+h_1}(w-M) + \dfrac{h_1}{h_0+h_1}(\hat{e}_1(h_1)-e_1) > \mu. \end{cases} \tag{36}$$

Let us note that:

$$\begin{aligned} \mu\eta + \frac{h_0}{h_0+h_1}w + \frac{h_1}{h_0+h_1}(\hat{e}_1(h_1)-e_1) < 0 \\ \implies \qquad \frac{h_0}{h_0+h_1}(w-M) + \frac{h_1}{h_0+h_1}(\hat{e}_1(h_1)-e_1) \leq 0, \end{aligned}$$

and that

$$\begin{aligned} \mu\eta + \frac{h_0}{h_0+h_1}(w-M) + \frac{h_1}{h_0+h_1}(\hat{e}_1(h_1)-e_1) > \mu \\ \implies \qquad \frac{h_0}{h_0+h_1}w + \frac{h_1}{h_0+h_1}(\hat{e}_1(h_1)-e_1) \geq 0. \end{aligned}$$

Moreover, $h_2^-(h_1)$ is interior if and only if

$$\frac{h_0}{h_0+h_1}w + \frac{h_1}{h_0+h_1}(\hat{e}_1(h_1)-e_1) \;\in\; [-\mu\eta, 0),$$

while $h_2^+(h_1)$ is interior if and only if

$$\frac{h_0}{h_0+h_1}w + \frac{h_1}{h_0+h_1}(\hat{e}_1(h_1)-e_1) \;\in\; \left(\frac{h_0}{h_0+h_1}M, \mu(1-\eta) + \frac{h_0}{h_0+h_1}M\right].$$

Consequently, with h_2^- and h_2^+ defined by (34)-(35)-(36), the agent's choice of period-2 activity, denoted by $h_2^\ddagger$, is given by

$$h_2^\ddagger(e_1,h_1) =$$

$$\begin{cases} 0 & \text{if } \dfrac{h_0}{h_0+h_1}w + \dfrac{h_1}{h_0+h_1}(\hat{e}_1(h_1)-e_1) \leq -\mu\eta, \\ h_2^-(e_1,h_1) & \text{if } \dfrac{h_0}{h_0+h_1}w + \dfrac{h_1}{h_0+h_1}(\hat{e}_1(h_1)-e_1) \in (-\mu\eta, 0), \\ h_2^*(h_1) & \text{if } \dfrac{h_0}{h_0+h_1}w + \dfrac{h_1}{h_0+h_1}(\hat{e}_1(h_1)-e_1) \in \left[0, \dfrac{h_0}{h_0+h_1}M\right], \\ h_2^+(e_1,h_1) & \text{if } \dfrac{h_0}{h_0+h_1}w + \dfrac{h_1}{h_0+h_1}(\hat{e}_1(h_1)-e_1) \in \left(\dfrac{h_0}{h_0+h_1}M, \mu(1-\eta) + \dfrac{h_0}{h_0+h_1}M\right), \\ +\infty & \text{if } \dfrac{h_0}{h_0+h_1}w + \dfrac{h_1}{h_0+h_1}(\hat{e}_1(h_1)-e_1) \geq \mu(1-\eta) + \dfrac{h_0}{h_0+h_1}M. \end{cases}$$

Hence, $h_2^\ddagger$ is continuous.

Remark: Whenever $\hat{w}(h_1,h_2) = w$, the solution to the agents' period-2 activity choice coincides with the solution $h^\dagger$ to (32)), i.e., absent headstart inequality.

*Choice of period-*1 *effort.* For a given choice of period-1 activity, the agent's period-1 effort $e_1^\ddagger$ is a solution, if any, to[77]

$$\begin{aligned}&\max_{e_1}\ \mathbb{E}_{\theta+\varepsilon_1}\Bigg[\frac{\mu h_1}{h_0+h_1+h_2^\ddagger}[\theta+\varepsilon_1+e_1+w-\hat{e}_1(h_1)-\hat{w}(h_1,h_2)]\\&\qquad\qquad+\frac{\mu h_2^\ddagger}{h_0+h_1+h_2^\ddagger}\frac{h_1}{h_0+h_1}[\theta+\varepsilon_1+w-\hat{w}(h_1,h_2)]\\&\qquad\qquad+\mu\eta\hat{e}_1(h_1)-\delta^{-1}g(e_1)+\mu\chi\hat{w}(h_1,h_2)+\mu\eta e_2^*(h_1,h_2^\ddagger)-g(e_2^*(h_1,h_2^\ddagger))\Bigg]\\&=\max_{e_1}\ \mathbb{E}_{\theta+\varepsilon_1}\Bigg[\frac{\mu h_1}{h_0+h_1+h_2^\ddagger}[e_1-\hat{e}_1(h_1)]-\delta^{-1}g(e_1)-\mu\Big(\frac{h_1}{h_0+h_1}-\chi\Big)\hat{w}(h_1,h_2)\\&\qquad\qquad\qquad\qquad+\mu\eta e_2^*(h_1,h_2^\ddagger)-g(e_2^*(h_1,h_2^\ddagger))\Bigg]\end{aligned}$$

This program has a solution $e_1^\ddagger<\infty$ as the objective is strictly decreasing for e_1 sufficiently high (above a finite threshold).[78]

Specifically, let us distinguish two cases.

(a) Whenever

$$\frac{h_0}{h_0+h_1}w+\frac{h_1}{h_0+h_1}(\hat{e}_1(h_1)-e_1)\in(-\mu\eta,0]\cup\Big[\frac{h_0}{h_0+h_1}M,\mu(1-\eta)+\frac{h_0}{h_0+h_1}M\Big),\quad(37)$$

period-2 precision $h_2^\ddagger(e_1,h_1)$ is given by the first-order conditions (34), which we can write as

$$\frac{\mu h_2^\ddagger}{h_0+h_1+h_2^\ddagger}=\mu\eta+\frac{h_0}{h_0+h_1}(w-\hat{w}(h_1,h_2^\ddagger))+\frac{h_1}{h_0+h_1}(\hat{e}_1(h_1)-e_1),\quad(38)$$

Hence, the program for period-1 effort is then strictly concave as its second derivative with respect to e_1 has the same sign as: $[h_1/(h_0+h_1)]^2-1<0$,[79] and the first-order condition

[77]Noting that

$$\frac{\mu h_2^\ddagger}{h_0+h_1+h_2^\ddagger}\frac{h_1}{h_0+h_1}=\Big(1-\frac{\mu(h_0+h_1)}{h_0+h_1+h_2^\ddagger}\Big)\frac{h_1}{h_0+h_1}=\frac{h_1}{h_0+h_1}-\frac{\mu h_1}{h_0+h_1+h_2^\ddagger}.$$

[78]This monotonicity stems from the cost of effort g being quadratic, and $h_2^\ddagger(h_1,e_1)=0$ for any e_1 sufficiently high.

[79]Indeed,

$$\frac{\mu(h_0+h_1)}{(h_0+h_1+h_2^\ddagger)^2}\frac{dh_2^\ddagger}{de_1}=-\frac{h_1}{h_0+h_1}$$

for an interior e_1 to be a solution writes as

$$\delta^{-1}e_1^\ddagger = \frac{\mu h_1}{h_0+h_1+h_2^\ddagger} + \frac{h_0 h_1}{(h_0+h_1)^2}(w-\hat{w}(h_1,h_2^\ddagger)) \tag{39}$$

As a consequence, whenever e_1 and $h_2^\ddagger$ are both given by their respective first-order conditions, then from (38),

$$\frac{\mu h_2^\ddagger}{h_0+h_1+h_2^\ddagger} + \frac{h_1}{h_0+h_1}e_1^\ddagger = \mu\eta + \frac{h_1}{h_0+h_1}\hat{e}_1(h_1) + \frac{h_0}{h_0+h_1}(w-\hat{w}(h_1,h_2^\ddagger))$$

which, using (39), can be rewritten as

$$\begin{aligned}&\frac{\mu h_2^\ddagger}{h_0+h_1+h_2^\ddagger} + \frac{h_1}{h_0+h_1}\frac{\delta\mu h_1}{h_0+h_1+h_2^\ddagger}\\ &\quad = \mu\eta + \frac{h_1}{h_0+h_1}\hat{e}_1(h_1) + \frac{h_0}{h_0+h_1}\left[1-\delta\left(\frac{h_1}{h_0+h_1}\right)^2\right](w-\hat{w}(h_1,h_2^\ddagger))\end{aligned}$$

Hence,

$$\frac{\mu h_2^\ddagger}{h_0+h_1+h_2^\ddagger} + \frac{h_1}{h_0+h_1}\frac{\delta\mu h_1}{h_0+h_1+h_2^\ddagger} \underset{(<)}{>} \mu\eta + \frac{h_1}{h_0+h_1}\hat{e}_1(h_1) \qquad \Longleftrightarrow \qquad w \underset{(<)}{>} \hat{w}(h_1,h_2^\ddagger).$$

In particular, as the LHS is a strictly increasing function of $h_2^\ddagger$, then for $\hat{e}(h_1) \equiv e_1^*(h_1) = \delta\mu h_1/(h_0+h_1+h_2^*(h_1))$,

$$h_2^\ddagger \underset{(<)}{>} h_2^*(h_1) \qquad \Longleftrightarrow \qquad w \underset{(<)}{>} \hat{w}(h_1,h_2^\ddagger),$$

and

$$h_2^\ddagger = h_2^*(h_1) \qquad \Longleftrightarrow \qquad w = \hat{w}(h_1,h_2^\ddagger).$$

Consequently,

$$e_1^\ddagger = e_1^*(h_1) \qquad \Longleftrightarrow \qquad w = \hat{w}(h_1,h_2^\ddagger).$$

By (39), a corner solution $e_1^\ddagger = 0$ may arise only for $w - \hat{w}(h_1,h_2^\ddagger) < 0$, hence $w = 0$ and $\hat{w}(h_1,h_2^\ddagger) = M$. A necessary condition is thus that

$$\frac{\mu h_1}{h_0+h_1+h_2^\ddagger} - \frac{h_0 h_1}{(h_0+h_1)^2}M \le 0$$

and thus by (38), that

$$\left(\mu - \frac{\mu h_0}{h_0 + h_1 + h_2^\ddagger} - \mu\eta + \frac{h_0}{h_0 + h_1}M - \frac{h_1}{h_0 + h_1}e_1^*(h_1)\right) - \frac{h_0 h_1}{(h_0 + h_1)^2}M \leq 0$$

which is equivalent to

$$\left[\mu(1-\eta) - \frac{\mu h_0}{h_0 + h_1 + h_2^\ddagger} - \frac{\delta h_1}{h_0 + h_1}\left(\mu(1-\eta) - \frac{\mu h_0}{h_0 + h_1 + h_2^*(h_1)}\right)\right] + \left(\frac{h_0}{h_0 + h_1}\right)^2 M \leq 0,$$

which is violated as $h_2^\ddagger < h_2^*(h_1)$ (since necessarily $w < \hat{w}(h_1, h_2^\ddagger$, as noted above).

Consequently, when case (a) applies and $\hat{e}_1(h_1) \equiv e_1^*(h_1)$, $e_1^\ddagger$ is always interior.

(b) Whenever

$$\frac{h_0}{h_0 + h_1}w + \frac{h_1}{h_0 + h_1}(\hat{e}_1(h_1) - e_1) \notin (-\mu\eta, 0] \cup \left[\frac{h_0}{h_0 + h_1}M, \mu(1-\eta) + \frac{h_0}{h_0 + h_1}M\right),$$

then $h_2^\ddagger$ locally does not depend on e_1, and the local first-order condition for e_1 is thus given by[80]

$$e_1^\ddagger = \frac{\delta\mu h_1}{h_0 + h_1 + h_2^\ddagger}$$

If $\frac{h_0}{h_0 + h_1}w + \frac{h_1}{h_0 + h_1}\hat{e}_1(h_1) \geq \mu(1-\eta) + \frac{h_0}{h_0 + h_1}M + \frac{h_1}{h_0 + h_1}e_1$, then $h_2^\ddagger(e_1, h_1) = +\infty$ and then the first-order condition yields that $e_1^\ddagger = 0$, which satisfies the initial inequality if and only if:

$$\frac{h_0}{h_0 + h_1}w + \frac{h_1}{h_0 + h_1}\hat{e}_1(h_1) \geq \mu(1-\eta) + \frac{h_0}{h_0 + h_1}M.$$

However, for $\hat{e}_1(h_1) \equiv e_1^*(h_1)$, the above inequality is violated,[81] and thus $e_1^\ddagger(h_1) > 0$ and $h_2^\ddagger(h_1, e_1^\ddagger(h_1)) < +\infty$.

If $\frac{h_0}{h_0 + h_1}w + \frac{h_1}{h_0 + h_1}\hat{e}_1(h_1) \leq -\mu\eta + \frac{h_1}{h_0 + h_1}e_1$, then $h_2^\ddagger(e_1, h_1) = 0$ and then the first-order condition yields that $e_1^\ddagger = \delta\mu h_1/(h_0 + h_1)$, which satisfies the initial inequality if and

[80] The objective is then locally concave as its second derivative with respect to e_1 is equal to $-g''(e_1) = -1 < 0$.

[81] Indeed,

$$\begin{aligned}\frac{h_1}{h_0 + h_1}\frac{\delta\mu h_1}{h_0 + h_1 + h_2^*(h_1)} &= \frac{\delta\mu h_1}{h_0 + h_1}\left(1 - \frac{h_0 + h_2^*(h_1)}{h_0 + h_1 + h_2^*(h_1)}\right) \\ &= \frac{\delta\mu h_1}{h_0 + h_1}\left(1 - \frac{h_0}{h_0 + h_1 + h_2^*(h_1)} - \eta\right)\end{aligned}$$

only if

$$\frac{h_0}{h_0+h_1}w+\frac{h_1}{h_0+h_1}\hat{e}_1(h_1)\leq -\mu\eta+\delta\mu\left(\frac{h_1}{h_0+h_1}\right)^2.$$

However, for $\hat{e}_1(h_1)\equiv e_1^*(h_1)$, the above inequality is violated,[82] and thus $e_1^{\ddagger}(h_1)<\delta\mu h_1/(h_0+h_1)$ and $h_2^{\ddagger}(h_1,e_1^{\ddagger}(h_1))>0$.

Lastly, if $\frac{h_0}{h_0+h_1}w+\frac{h_1}{h_0+h_1}(\hat{e}_1(h_1)-e_1)\in\left(0,\frac{h_0}{h_0+h_1}M\right)$, then $h_2^{\ddagger}(e_1,h_1)=h_2^*(h_1)$ and then the first-order conditions yields that $e_1^{\ddagger}=\delta\mu h_1/[h_0+h_1+h_2^*(h_1)]=e_1^*(h_1)$, so that the initial condition becomes

$$\frac{h_1}{h_0+h_1}e_1^*(h_1)<\frac{h_0}{h_0+h_1}w+\frac{h_1}{h_0+h_1}\hat{e}_1(h_1)<\frac{h_0}{h_0+h_1}M+\frac{h_1}{h_0+h_1}e_1^*(h_1),$$

which is violated whenever $\hat{e}_1(h_1)\equiv e_1^*(h_1)$.

Consequently, for $\hat{e}_1(h_1)\equiv e_1^*(h_1)$, case (b) never arises, i.e. $h_2^{\ddagger}(e_1,h_1)$ is given by the first-order condition (38) and $e_1^{\dagger}$ by the first-order condition (39).

Hence, suppose that beliefs are such that $\hat{e}_1(h_1)\equiv e_1^*(h_1)$ for any $h_1\geq 0$. Then, from the above discussion, $\hat{w}(h_1,h_2)\equiv 0$ for any $h_2<h_2^*(h_1)$, and $\hat{w}(h_1,h_2)\equiv M$ for any $h_2>h_2^*(h_1)$, are consistent with our equilibrium concept.[83]. Therefore, by concavity, after any given choice of period-1 activity h_1 and period-1 effort e_1, the agent's period-2 activity choice is given by

$$h_2^{\ddagger}(h_1,e_1)=h_2^*(h_1).$$

As a consequence, for any period-1 activity choice, the agent's period-1 effort is given by $e_1^*(h_1)$.[84]

*Choice of period-*1 *activity.* Suppose that the audience's beliefs about period-1 effort are such that $\hat{e}_1(h_1)\equiv e_1^*(h_1)$ for any $h_1\geq 0$, while its beliefs about headstart are such that $\hat{w}(h_1,h_2)\equiv 0$ for any $h_2<h_2^*(h_1)$, and $\hat{w}(h_1,h_2)\equiv M$ for any $h_2>h_2^*(h_1)$. Then, the

[82]Indeed,

$$\begin{aligned}\frac{h_1}{h_0+h_1}\frac{\delta\mu h_1}{h_0+h_1+h_2^*(h_1)}&=\frac{\delta\mu h_1}{h_0+h_1}\left(1-\frac{h_0+h_2^*(h_1)}{h_0+h_1+h_2^*(h_1)}\right)\\&=\frac{\delta\mu h_1}{h_0+h_1}\left(\frac{h_1+h_2^*(h_1)}{h_0+h_1+h_2^*(h_1)}-\eta\right),\end{aligned}$$

and for any $h_2\geq 0$ (and in particular $h_2^*(h_1)$),

$$\frac{h_1+h_2}{h_0+h_1+h_2}>\frac{h_1}{h_0+h_1},$$

as the LHS strictly increases with h_2.

[83]In fact, they are the only monotone beliefs consistent with our equilibrium concept.

[84]The arguments are those used in case (b) above, in the subcase in which $h_2^{\ddagger}(h_1,e_1)=h_2^*(h_1)$.

agent's choice of h_1 is given by the solution, if any, to

$$\begin{aligned}
&\max_{h_1}\Bigg[\mu\eta\hat{e}_1(h_1)-\delta^{-1}g(e_1^{\ddagger}(h_1))+\mu\eta e_2^*(h_1,h_2^{\ddagger}(h_1,e_1^{\ddagger}(h_1)))-g(e_2^*(h_1,h_2^{\ddagger}(h_1,e_1^{\ddagger}(h_1)))\\
&\qquad+\frac{\mu h_1}{h_0+h_1+h_2^{\ddagger}(h_1,e_1^{\ddagger}(h_1))}[e_1^{\ddagger}(h_1)-\hat{e}_1(h_1)]+\frac{\mu(h_1+h_2^{\ddagger}(h_1,e_1^{\ddagger}(h_1)))}{h_0+h_1+h_2^{\ddagger}(h_1,e_1^{\ddagger}(h_1))}w\\
&\qquad-\mu\Big(\frac{h_1+h_2^{\ddagger}(h_1,e_1^{\ddagger}(h_1))}{h_0+h_1+h_2^{\ddagger}(h_1,e_1^{\ddagger}(h_1))}-\chi\Big)\hat{w}(h_1,h_2^{\ddagger})\Bigg]\\
=&\max_{h_1}\Bigg[\mu\eta e_1^*(h_1)-\delta^{-1}g(e_1^*(h_1))+\mu\eta e_2^*(h_1,h_2^*(h_1))-g(e_2^*(h_1,h_2^*(h_1)))\\
&\qquad+\frac{\mu(h_1+h_2^*(h_1))}{h_0+h_1+h_2^*(h_1)}w-\mu\Big(\frac{h_1+h_2^*(h_1)}{h_0+h_1+h_2^*(h_1)}-\chi\Big)\hat{w}(h_1,h_2^*(h_1))\Bigg].
\end{aligned}$$

Since $\mu\eta e_2^*(h_1,h_2^*(h_1))-g(e_2^*(h_1,h_2^*(h_1)))$ does not depend on h_1, the agent's choice of h_1 is given by the solution to:

$$\begin{aligned}
&\max_{h_1}\Bigg[\mu\eta e_1^*(h_1)-\delta^{-1}g(e_1^*(h_1))\\
&\qquad+\frac{\mu(h_1+h_2^*(h_1))}{h_0+h_1+h_2^*(h_1)}w-\mu\Big(\frac{h_1+h_2^*(h_1)}{h_0+h_1+h_2^*(h_1)}-\chi\Big)\hat{w}(h_1,h_2^*(h_1))\Bigg]
\end{aligned}\tag{40}$$

Equilibrium existence. Let us look for dynamic versions of distinction and displacement. *Distinction.* Suppose $2\eta<\chi$. Let $h_P=h_1^*$ and $h_R>h_1^*$ be given by

$$\begin{aligned}
&\mu\eta e_1^*(h_1^*)-\delta^{-1}g(e_1^*(h_1^*))+\mu\eta e_2^*(h_1^*,h_2^*(h_1^*))-g(e_2^*(h_1^*,h_2^*(h_1^*)))\\
&\quad=\mu\eta e_1^*(h_R)-\delta^{-1}g(e_1^*(h_R))\\
&\qquad+\mu\eta e_2^*(h_R,h_2^*(h_R))-g(e_2^*(h_R,h_2^*(h_R))\\
&\qquad\qquad-\mu\Big(\frac{h_R+h_2^*(h_R)}{h_0+h_R+h_2^*(h_R)}-\chi\Big)M,
\end{aligned}$$

i.e. equivalently, as $\mu\eta e_2^*(h_1,h_2^*(h_1))-g(e_2^*(h_1,h_2^*(h_1)))$ does not depend on h_1, by

$$\mu\eta e_1^*(h_1^*)-\delta^{-1}g(e_1^*(h_1^*))\;=\;\mu\eta e_1^*(h_R)-\delta^{-1}g(e_1^{\ddagger,P}(h_R))-\mu\Big(\frac{h_R+h_2^*(h_R)}{h_0+h_R+h_2^*(h_R)}-\chi\Big)M.$$

Let the audience's beliefs about head starts be given by

$$\hat{w}(h_1,h_2^*(h_1))=\begin{cases}0 & \text{for any } h_1<h_R,\\ M & \text{for any } h_1\geq h_R.\end{cases}$$

and for any $h_2 \neq h_2^*(h_1)$,[85]

$$\hat{w}(h_1, h_2) = \begin{cases} 0 & \text{if } \; h_2 < h_2^*(h_1), \\ M & \text{if } \; h_2 > h_2^*(h_1). \end{cases}$$

and its beliefs about effort be $\hat{e}_1(h_1) = e_1^*(h_1, h_2^*(h_1))$, $\hat{e}_2(h_1, h_2) = e_2^*(h_1, h_2)$.

Hence, from our above discussion, for any $h_1 \geq 0$, $h_2^{\ddagger,P} = h_2^{\ddagger,R} = h_2^*(h_1)$, and $e_1^\ddagger = e_1^*(h_1)$.

Let us check that these strategies and beliefs form a (fully separating) equilibrium. By concavity, for all $h > h_R$,

$$\begin{aligned} &\mu\eta e_1^*(h_R) - \delta^{-1} g(e_1^*(h_R)) + \mu\eta e_2^*(h_R, h_2^*(h_R)) - g(e_2^*(h_R, h_2^*(h_R))) \\ &\quad \geq \; \mu\eta e_1^*(h) - \delta^{-1} g(e_1^*(h)) + \mu\eta e_2^*(h, h_2^*(h)) - g(e_2^*(h, h_2^*(h))), \end{aligned}$$

and for all $h < h_P$,

$$\begin{aligned} &\mu\eta e_1^*(h_P) - \delta^{-1} g(e_1^*(h_P)) + \mu\eta e_2^*(h_P, h_2^*(h_P)) - g(e_2^*(h_P, h_2^*(h_P))) \\ &\quad \geq \; \mu\eta e_1^*(h) - \delta^{-1} g(e_1^*(h)) + \mu\eta e_2^*(h, h_2^*(h)) - g(e_2^*(h, h_2^*(h))). \end{aligned}$$

Therefore, the above strategies and beliefs form a (fully separating) equilibrium if and only if the two following conditions hold: for all $h < h_R$,

$$\begin{aligned} &\mu\eta e_1^*(h_R) - \delta^{-1} g(e_1^*(h_R)) + \mu\eta e_2^*(h_R, h_2^*(h_R)) - g(e_2^*(h_R, h_2^*(h_R))) \\ &\quad \geq \; \mu\eta e_1^*(h) - \delta^{-1} g(e_1^*(h)) + \mu\eta e_2^*(h, h_2^*(h)) - g(e_2^*(h, h_2^*(h)) + \mu\Big(\frac{h + h_2^*(h)}{h_0 + h + h_2^*(h)} - \chi\Big)M \end{aligned}$$

and that for all $h > h_R$,

$$\begin{aligned} &\mu\eta e_1^*(h_P) - \delta^{-1} g(e_1^*(h_P)) + \mu\eta e_2^*(h_P, h_2^*(h_P)) - g(e_2^*(h_P, h_2^*(h_P))) \\ &\quad \geq \; \mu\eta e_1^*(h) - \delta^{-1} g(e_1^*(h)) + \mu\eta e_2^*(h, h_2^*(h)) - g(e_2^*(h, h_2^*(h)) - \mu\Big(\frac{h + h_2^*(h)}{h_0 + h + h_2^*(h)} - \chi\Big)M \end{aligned}$$

The same arguments as in the proof of Proposition 1 (see Appendix C) then yield the result.

[85]The condition $h_2 < h_2^*(h_1)$ (resp. $>$) is equivalent to

$$\frac{\mu h_2}{h_0 + h_1 + h_2} + \frac{\delta \mu h_1}{h_0 + h_1 + h_2} < \mu\eta + \frac{\delta \mu h_1}{h_0 + h_1 + h_2^*(h_1)}, \qquad \text{(resp. } >\text{)}.$$

Displacement. Suppose $2\eta > \chi$. Let $h_R = h_1^*$ and $h_P < h_1^*$ be given by

$$\begin{aligned}
&\mu\eta e_1^*(h_P) - \delta^{-1} g(e_1^*(h_P)) + \mu\eta e_2^*(h_1^*, h_2^*(h_1^*)) - g(e_2^*(h_1^*, h_2^*(h_1^*))) \\
&\quad = \mu\eta e_1^*(h_1^*) - \delta^{-1} g(e_1^*(h_1^*)) \\
&\qquad + \mu\eta e_2^*(h_1^*, h_2^*(h_1^*)) - g(e_2^*(h_1^*, h_2^*(h_1^*)) \\
&\qquad\qquad - \mu\Big(\frac{h_1^* + h_2^*(h_1^*)}{h_0 + h_1^* + h_2^*(h_1^*)} - \chi\Big)M,
\end{aligned}$$

i.e. equivalently, as $\mu\eta e_2^*(h_1, h_2^*(h_1)) - g(e_2^*(h_1, h_2^*(h_1)))$ does not depend on h_1, by

$$\mu\eta e_1^*(h_P) - \delta^{-1} g(e_1^*(h_P)) = \mu\eta e_1^*(h_1^*) - \delta^{-1} g(e_1^*(h_1^*)) - \mu\Big(\frac{h_1^* + h_2^*(h_1^*)}{h_0 + h_1^* + h_2^*(h_1^*)} - \chi\Big)M,$$

Let the audience's beliefs about head starts be given by

$$\hat{w}(h_1, h_2^*(h_1)) = \begin{cases} 0 & \text{for any } h_1 \leq h_P, \\ M & \text{for any } h_1 > h_P. \end{cases}$$

and for any $h_2 \neq h_2^*(h_1)$,[86]

$$\hat{w}(h_1, h_2) = \begin{cases} 0 & \text{if } \ h_2 < h_2^*(h_1), \\ M & \text{if } \ h_2 > h_2^*(h_1). \end{cases}$$

and its beliefs about effort be $\hat{e}_1(h_1) = e_1^*(h_1, h_2^*(h_1))$, $\hat{e}_2(h_1, h_2) = e_2^*(h_1, h_2)$.

The same arguments as in the proof of Proposition 1 (see Appendix C) again yield the result.

K Proofs of Proposition 8 and Corollary 3

We look for separating equilibria with degenerate off-path beliefs. Hence, with relative image concerns, by linearity, conditional on choosing an activity with precision h, an agent with wealth w (still) exerts effort $e^*(h)$ such that

$$g'(e^*(h)) = \frac{\mu h}{h_0 + h},$$

[86] The condition $h_2 < h_2^*(h_1)$ (resp. $>$) is equivalent to

$$\frac{\mu h_2}{h_0 + h_1 + h_2} + \frac{\delta\mu h_1}{h_0 + h_1 + h_2} < \mu\eta + \frac{\delta\mu h_1}{h_0 + h_1 + h_2^*(h_1)}, \qquad \text{(resp. } >\text{)}.$$

as the weights on within- and across-activity images sum to 1.

For any $h \in \mathbb{R}_+$, let $U(h,\zeta) \equiv \beta(e^*(h)) + \zeta\mu\eta e^*(h) - g(e^*(h))$. With relative image concerns, in a separating equilibrium with degenerate off-path beliefs, each agent with head start w chooses their activity by solving:

$$U(h,\zeta) + \frac{\mu h}{h_0 + h} w - \frac{\mu h}{h_0 + h}\mathbb{E}[w|h] + \zeta\mu\chi\mathbb{E}[w|h]$$

i.e. by solving (P), only replacing η by $\zeta\eta$ and χ by $\zeta\chi$. Proposition 8 and Corollary 3 then follow from the proofs of Proposition 1 (see Appendix C) and comparative statics with respect to ζ.