Meritocracy and Inequality

Paul-Henri Moisson*

September 11, 2023

PRELIMINARY & INCOMPLETE.
PLEASE DO NOT CIRCULATE.

Abstract

How do individuals behave in a society that vows to reward "merit", despite not all individuals being on the same starting line? Does inequality in headstarts make meritocracy undesirable? To answer these questions, this paper develops a model of career concerns in which agents publicly choose between several activities in which to exert effort, and differ along a privately observable characteristic ("headstart") that affects their performance. The agents' audience values their talent, effort and headstart. We highlight the race between two effects: a displacement effect by which the "poor" (headstart-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 headstart 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. We interpret the model in terms of "meritocracy", characterize optimal activity design and discuss several policy implications, emphasizing the unintended consequences of common interventions.

JEL numbers: D2, D6, H2, J24.

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

*Toulouse School of Economics, 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.[1] The meritocratic vocabulary has been increasingly popular since it was first coined by Young (1958), and "meritocracy" itself increasingly perceived as a desirable aim for societies and organizations (again, with varying definitions of merit).[2] Yet, meritocracy has also faced rising and multifaceted criticisms (e.g. Markovits 2019, Sandel 2020). On the one hand, it is argued that "perfect meritocracy" cannot be achieved as any attempt at implementing it is rigged by inequalities in individuals' "headstarts" – which encompass not only financial wealth, but also human and social capital more broadly. Critics have complained that while pretending to reward "merit", meritocracy is in fact rewarding headstarts, and that these have low social value – e.g. private lessons aimed at securing a high score at an exam, but not improving the student's long-term productivity. The counterargument claims that, even ignoring headstarts, the headstart-rich deserve their higher status as they exert more effort than the headstart-poor – e.g. by pursuing longer and more demanding degrees, or putting up longer working hours.[3] Hence, on the other hand, going back to Young's (1958) original stance, it is argued that even a "perfect meritocracy" would not be desirable because of the inequality it induces, the social stigma it inflicts on the "losers", and, as emphasized more recently, the excessive competition it induces among the "winners". Could a unified model reconcile these seemingly contradicting claims? More specifically, we ask three guiding questions: What is the interplay of meritocracy and headstart inequality? For a given headstart inequality, what would a "second-best" meritocracy look like? What would the policy implications be?

We thus study an environment in which agents care about their "merit", as perceived by others (peers, future employers, universities, society, etc.). We follow a positive approach, letting "merit" be a given combination of talent and effort, with varying weights. We augment the canonical career-concerns model (Holmström 1982/1999) by adding three key features. Firstly, we allow agents to publicly choose among several activities – e.g. majors at school, or colleges, or jobs, or tasks within an organization, etc. Secondly, we introduce a privately observable heterogeneity among agents that affects their (public) performance. We refer to

[1]This paper does not argue about what the right definition of merit is, but rather considers a wide range of possible definitions and investigate their consequences.

[2]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. Part of the seduction of the meritocratic vocabulary may lie in the idea that it "rewards" something, i.e. that under certain (possibly vague or unreachable) conditions, individuals *deserve* a reward.

[3]Underlining the psychological strength of this argument, recent experimental evidence seems to suggest that when rewarding effort (allocating "merit"), individuals do not fully take into account (if at all) the incentives that the agents had to exert effort independently of their "merit" expectations – e.g. their monetary incentives. See notably Andre (2022).

this heterogeneous trait as "headstart", whether it stems from financial resources, human capital, social capital, etc. Thirdly, we allow (expected) effort and (expected) headstart to be valued by the agents' audience – e.g. if effort has a longlasting impact on an agent's productivity, or if headstarts embody soft skills, cultural capital, intrinsic motivation to perform or positive externalities from headstart-rich peers. As a consequence, agents care not only about their talent image, but also about their effort and headstart images.

We show that headstart inequality generates separating equilibria and that, under a standard equilibrium concept (Bayesian perfection with D1), their structure can be explicitly characterized. Indeed, the audience's weights on an agent's individual talent on the one hand, and on her headstart (or activity peers' headstart) on the other, induce contrasting incentives. When talent image concerns dominate headstart image ones, a *displacement* effect arises: the "poor" (headstart-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 performance-based rewards. Hence, the rich choose an activity with high rewards on performance and displace the poor towards less rewarding 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) headstart.

Yet, headstart inequality has another facet. When headstart 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 performance –, thereby reaping a higher headstart image, while foregoing the higher talent image they would obtain by pooling with the poor. As a consequence, higher headstart inequality reduces aggregate effort if the displacement effect dominates, and increases it if the distinction effect does. In both cases, headstart inequality drives the rich and the poor apart.[4]

Lastly, the agents' choice of precision depends on the payoff from expected effort, including both the weight on the effort image (and thus possibly future wages) and current-period transfers if based on expected current effort. All else being equal, the higher the payoff from expected effort, the higher the agents' favored precision.

[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 headstart 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 (and remaining) label for the opposed effect.

In a few words, our three additions to the standard career-concerns model combine as follows: headstart inequality makes agents willing to separate, multiple activities allow them to do so, while the respective weights on talent, effort and headstart images determine in which direction agents separate.

Concretely, in the context of education, when the displacement effect dominates, students from disadvantaged backgrounds fret about confronting well-prepared and/or well-connected 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 and/or well-connected students engage in highly selective and demanding tracks to discourage disadvantaged students from following them, and thus enjoy the collective reputation attached to their preparation and/or connections. Displacement prevails when image weights emphasize talent over headstart – e.g. when scientific intuition is more highly valued than exam preparation thanks to private lessons –, whereas distinction prevails when headstarts are praised – e.g. when (certain) manners, soft skills or social connections are highly valued by recruiters.

"Second-best meritocracy". Hence, for a given headstart inequality, what would the "second-best meritocracy" look like?[5] We focus on separating equilibria for our policy analysis. We begin by studying the optimal activity characteristics (precision and transfers). Namely, 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", which is beyond the principal's control – e.g. in the context of (national) education, drop out of the schooling system, or leave 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 reduce the rents of the party less tempted to leave. The magnitude of displacement and distinction increases with headstart inequality. Hence, for sufficiently low headstart inequality, the principal chooses the former solution, designing "aligned incentives" (such that the poor are the most tempted to leave) if the precision of the outside activity is below the first-best precision, and "countervailing incentives" (such that the rich are the most tempted to leave) otherwise. However, for sufficiently large headstart inequality, the principal chooses the latter solution, designing aligned incentives if distinction prevails in the outside activity (which reduces the rich's rent), and countervailing incentives if on the opposite, displace-

[5]The "first-best meritocracy" would correspond to the case in wich headstarts would be public.

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

ment prevails in the outside activity (which reduces the poor's rent). As a consequence, for sufficiently large headstart inequality, the principal may choose the same precision for both the rich and the poor.

Common-policy recommendations. Turning toward more limited interventions (sill focusing on separating equilibria), we study several policies and highlight their unintended consequences stemming from the displacement and distinction effects. One such popular intervention is capping activity precision.[7] When distinction dominates under *laissez-faire*, a cap on activity precision destroys all separating equilibria as it "corners" the rich in a lower-precision activity, closer to the poor's *laissez-faire* activity, and thereby helps the poor to join them. By contrast, when displacement dominates under *laissez-faire*, separating equilibria still exist with a cap on activity precision: 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 need to move toward an activity with even lower precision to avoid the rich. As a consequence, when displacement dominates under *laissez-faire*, both the poor 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 headstart inequality can create a "whack-a-mole" game for the policy maker, in which one party (rich or poor) chases the other while circumventing any policy intervention.

We then investigate income taxation. In our risk-neutral setting, headstart equality implies that the optimal income tax is nil.[8]. By contrast, with headstart inequality, the optimal income tax depends crucially on whether displacement or distinction dominates. Indeed, with displacement (resp. distinction), headstart 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 headstart 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 comparative statics of the optimal intensity of image concerns with respect to headstart inequality.[10] The

[7]In France, "selection" in public universities has been opposed by many political leaders and intellectuals. In the United States, Sandel's (2020) proposal of a lottery among qualified students for admission to elite colleges can also be interpreted as an attempt 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 headstart inequality when displacement prevails can be interpreted as another illustration of the "whack-a-mole" policy game induced by image concerns and headstart inequality. In fact, this result is related to Rothschild and Scheuer's (2016) analysis of optimal taxation with rent-seeking. In our environment, rents stem from privately observable headstart which affects the agents' performance and image, and the rent-seeking unfolds across activities.

[10]A higher intensity of image concerns may stem for instance from a heightened visilibity or emphasis on "merit", or a steeper slope for its associated consequences (be they material rewards or social status). The

higher the headstart inequality, the lower the optimal intensity of image concerns if distinction prevails, but the higher the optimal intensity if displacement does.

Lastly, we consider two important extensions. Introducing a second period – e.g. college after high school, or grad school after undergrad – to capture (some) dynamics of our model yields that there (generically) exists separating equilibria in which the rich and the poor not only separate in their first period activity choices, thereby revealing their headstart, but also in the second period, with the rich again choosing 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' utility. 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). In addition, relative image concerns deliver several interesting insights. In particular, the more a society is segregated along activity lines – i.e. the more individuals compare their payoffs only to their activity peers –, the larger the magnitude of the displacement effect and the lower the one of the distinction effect.

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) headstarts and investigating the consequences of headstart inequality. It unveils the basic mechanisms and key drivers of the following analyses. Section 3 studies optimal activity characteristics and policy. Section 4 considers several extensions. Section 5 reviews the literature. Section 6 concludes by briefly evoking several alleys for future research. All proofs are in the Appendix.

2 Model (laissez-faire)

2.1 No headstart inequality (Homogeneous agents)

There is a continuum of agents, with mass 1. Each agent participates in exactly one among N activities, indexed by $k \in \{1, ..., N\}$. Agents may for instance be students choosing a major, or a college, or whether doing a PhD (and in which field), or they could be

trade-off between incentivizing effort and distorting activity choices is related to the trade-off for optimal privacy between incentivizing effort and learning societal preferences, described by Ali and Bénabou (2020).

prospective workers choosing among job offers from different firms or industries. Each agent is characterized by her (unobservable) talent $\theta \in \mathbb{R}$.

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

$$y_k = \theta + e_k + \varepsilon_k,$$

where $e_k \geq 0$ is the agent's effort in activity k and ε_k a random noise, normally distributed with mean zero and precision h_k (inverse of variance).

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

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

Career/Image concerns. An agent's *audience* may be thought of as embodying (non-exclusively) the other agents, third-parties such as relatives and friends, the agent's supervisors or managers, potential future employers, etc. Each agent values her audience's opinion about herself. Namely, the agent cares about the audience's posterior belief on the weighted sum of her type and, unlike in Holmström (1982), her effort given her choice of activity k and outcome y_k:[11]

$$\psi(y_k) \equiv \hat{\theta}(y_k) + \eta \hat{e}(y_k)$$

with $\eta > 0$ the weight on the effort image, capturing in particular the long-lasting impact of effort on productivity – e.g. θ may be interpreted as an agent's innate ability, and the product ηe as her acquired ability.[12]

In our main utility specification, agents have *absolute* image concerns, i.e. each agent cares about her image $\psi(y_k)$, regardless of other agents' own images.[13] Complementarily, we consider in Section 4.2 *relative* image concerns, with each agent caring about her image relative to the images of some reference groups (in the spirit of Merton 1957) to capture, e.g.

[11] For conciseness, we commit a slight abuse of notation as all images depend both on the activity choice k and the performance y_k.

[12] All our results go through if $\eta \leq 0$, e.g. if effort today damages one's future productivity (say, due to harmful activities). Our main insights are also robust to allowing η to vary across activities.

[13] We assume that the choice of activity does not signal anything about θ – this "no-signaling-what-you-don't-know" property is implied by PBE (Fudenberg-Tirole 1991).

the positional property of social status, the distinction between specialized and generalist skills, or contest-like features (say, in university admission).[14] Importantly and maybe surprisingly, all our insights, both positive and normative, hold whether agents have absolute or relative image concerns (see Section 4.2).

In addition to her image concerns, each agent receives a direct transfer $\beta_k \in \mathbb{R}$ (fixed wage or tuition fee), which depends on the activity k chosen by the agent. [We take the transfers $(\beta_k)_k$ as exogenous in this Section, but we will endogenize them in Sections 2.3 and 3.] Therefore, each activity $k \in \{1, ..., N\}$ is characterized by its precision h_k and its transfer β_k.

Each agent thus chooses her activity k and her effort levels (e_k) to solve:

$$\max_{k \in \{1,...,N\}} \max_{e_k \geq 0} \mathbb{E}\Big[\beta_k + \mu\psi(y_k) - g(e_k)\Big],$$

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

The equilibrium concept is Bayesian Perfection.

Preliminary analysis. The audience updates its belief given the agent's activity choice k and performance y_k:

$$\hat{\theta}(y_k) + \eta\hat{e}(y_k) = \frac{h_k}{h_0 + h_k}(y_k - e_k^*) + \eta e_k^*.$$

An agent's optimal effort in activity k, denoted by e_k^* is given by:

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

In particular, all agents exert the same effort in a given activity.

Letting for all $k \in \{1, ..., N\}$,

$$U_k \equiv \mu\eta e_k^* + \beta_k - g(e_k^*),$$

[14] As an illustration, when an agent's single reference group is the whole population, relative image concerns write as

$$\psi(y_k) - \mathbb{E}[\psi(y)],$$

where the expectation is taken over all activities, given the anticipated characteristics and effort levels of agents choosing each activity.

each agent thus chooses her activity by solving:

$$\max_{k\in\{1,...,N\}} U_k.$$

We assume that this maximization program has a unique solution, a generic property. Hence, in equilibrium, all agents choose the same activity and the same effort level.

Lemma 1 (**Homogeneous agents**). *An agent's equilibrium effort in a given activity k increases with the intensity of image concerns μ and the precision h_k of the activity, but does not depend on the long-lasting productivity of effort η. All else being equal, a higher productivity of effort induces the agents to choose an activity with a higher precision.*

The agents' taste for precision h depends on the long-lasting productivity of effort η. The higher the long-lasting productivity of effort η, the higher the precision favored by the agents. Indeed, the higher the long-lasting productivity of effort η, the larger the weight on effort in the agents' images, i.e. the more the agents care about the audience's expectation of their effort, and the latter increases with higher precision (or more generally higher-powered incentives).

2.2 Headstart inequality (Heterogeneous agents)

Suppose now that some agents have a headstart but others don't, and that while agents privately know whether they enjoy a headstart, the audience does not observe it. We refer to these headstarts as "headstart" in the largest meaning, i.e. encompassing not only financial headstart, but also human and/or social capital – e.g. soft skills, social connections or even intrinsic motivation to perform.

For simplicity, suppose that each agent has a "headstart" $w \in \{0, M\}$. Hence, M is a measure of headstart inequality.[15,16] Headstarts are i.i.d. across agents, and let $p \equiv \mathbb{E}[w]/M$ denote the share of the rich in the population. An agent's outcome in activity k when having headstart w and exerting effort e_k now writes as[17]

$$y_k = \theta + e_k + w + \varepsilon_k,$$

[15] Our main insights obtain with only two levels of headstart, as only headstart differences matter.

[16] Importantly, our results still hold as long as one's headstart is at least *imperfectly* observable.

[17] For simplicity, we (tacitly) assume throughout the exposition that each agent always spends all her headstart. A rationale is that the agents' (marginal) disutility from spending their headstart is sufficiently low. This is highly likely to be the case in particular for "headstarts" stemming from human capital or soft skills, on which we focus in the exposition. More generally, each agent could incur a disutility from spending her headstart, so that spending $m \leq w$ entails a cost $g_w(m)$, with $g_w(\cdot) : \mathbb{R}_+ \to \mathbb{R}_+$, increasing and convex. Our main insights would remain.

The audience is interested in the weighted sum of an agent's (individual) talent, effort and headstart (either own or activity peers'). An agent's image is given by:

$$\psi(y_k) \equiv \hat{\theta}(y_k) + \eta\hat{e}(y_k) + \chi\hat{m}(y_k),$$

with $\chi \geq 0$ the weight on headstart image, capturing how long-lasting and productive/socially useful private investments are. The difference $1-\chi$ may thus be interpreted as a measure of unproductive signalling or as the difference between the short-term and the long(er)-term productive impacts of private investments – the analogue for effort being the difference $1-\eta$. The parameters χ and η may also capture peer externalities (in a reduced form): χ may thus be interpreted as the marginal value of having headstart-rich (activity) peers, and η as the marginal value from peers' effort.[18]

We focus on Bayesian perfect equilibria that satisfy the D1 criterion.[19] We look for separating equilibria.

Preliminary analysis. Because effort and private investment enter additively in an agent's performance, an agent's optimal effort level in a given activity does not depend on her headstart and is still given by (1). Yet, an agent's activity choice now depends on her headstart and on the other agents' activity choices.

Namely, in a separating equilibrium, each agent with headstart w chooses her activity by solving:

$$\max_{k\in\{1,\ldots,N\}} \left(U_k + \frac{\mu h_k}{h_0+h_k}w - \mu\Big(\frac{h_k}{h_0+h_k} - \chi\Big)\mathbb{E}[w|k]\right) \tag{P}$$

Hence, with headstart inequality and under separation, an agent's choice of activity has three drivers:

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

(ii) an incentive stemming from the private benefits of her own headstart w, that accrue via her (boosted) talent image,

(iii) an incentive stemming from the audience's expectation of her headstart and/or the

[18] The weights on effort and headstart images, η and χ are relative to the weight on talent image, normalized to 1.

[19] Our refinement is *in the spirit of* D1 (as defined by Banks and Sobel 1987). We refer to Appendix C for a precise definition. In words, we require that if there exist two agent types t, t' such that whenever a type-t agent either wishes to deviate to an off-path activity k or is indifferent (for some induced beliefs), a type-t' agent stricly wishes to deviate to that activity (for the same induced beliefs), then the audience puts zero probability on an agent in activity k having type t.

collective impact of her activity peers' headstart, $\mathbb{E}[w|k]$.

As we shall see, the second and third term in (P) drive the agents' willingness to separate, and the second term generates the sorting condition, determining how they separate.

Indeed, let us consider the third term. Whenever some rich choose the same activity as some poor, the rich's headstart inflicts on the poor a negative externality on their (individual) talent image, while conversely, the poor's presence confers a positive externality on the rich's talent image. The magnitude of these externalities increases with the activity's precision (proportionally to the belief updating coefficient $h_k/(h_0 + h_k)$). Yet, the rich (resp. the poor) also bring a positive (resp. negative) externality on the poor's (resp. rich's) headstart image, with a magnitude proportional to the weight on headstart image, χ. As a consequence, when $h_k/(h_0 + h_k) > \chi$, i.e. when talent image concerns dominate headstart ones, the rich are eager to blend with the poor (e.g. eager to take the same tests, engage in the same activities to boost their talent image), but the poor are eager to separate from them (to avoid the rich's competition and to safeguard their talent image). The opposite holds when $h_k/(h_0 + h_k) < \chi$, i.e. when headstart image concerns dominate talent ones: the poor are then eager to blend with the rich (e.g. to reap the benefits of having peers with strong soft skills or high social capital) while the rich are eager to separate from them (to signal to the audience that they are the ones with the strong softskills or high social capital). Hence, whenever $h_k/(h_0 + h_k) \neq \chi$, 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 is related to talent images. Indeed, as an agent's own headstart w improves her talent image, it is a complement to precision h. This induces the following sorting condition.

Lemma 2 (**Sorting condition with headstart inequality**). *Consider a given activity k and fix posterior belief* $\mathbb{E}[w|k] \in \{0, M\}$. *Let*

$$V_k(w) \equiv \left(U_k + \frac{\mu h_k}{h_0 + h_k} w - \mu\left(\frac{h_k}{h_0 + h_k} - \chi\right)\mathbb{E}[w|k]\right).$$

Then, within activity k, a rich agent prefers a strictly higher precision h_k than a poor agent:

$$\frac{\partial}{\partial h_k}\Big(V_k(M) - V_k(0)\Big) > 0.$$

The sorting condition thus yields that the poor separate from the rich by moving toward activities with lower precision, in which the rich's headstart is less effective, while on the opposite, the rich separate from the poor by moving toward activities with higher precision, in which their headstart is more painful to the poor.

As a consequence, in any separating equilibrium, the poor choose activities with lower precision (more generally, lower incentives on effort) than the rich.[20]

To derive a sharp description of separating equilibria for any headstart inequality $M \geq 0$, we consider the limit case of a continuum of activities.[21] For future reference, we single out our assumptions for this case.

Assumption 1 (Continuum of activities). *Suppose there exists a continuum of activities indexed by their precision $h \in \mathbb{R}_+$. We denote the transfer in activity h as $\beta(e^*(h))$, with $g'(e^*(h)) = \mu h/(h_0+h)$, i.e. as a function of the expected effort in activity h (equivalently as a function of precision h), which we assume to be continuous. Defining for any h, $U(h) \equiv \beta(e^*(h)) + \mu\eta e^*(h) - g(e^*(h))$, suppose that $U(h)$ is continuously differentiable, first strictly increasing then strictly decreasing, with a unique interior maximum. Let h^* be such that $U(h^*) = \max_{h \geq 0} U(h)$.*

The shape of $U(h)$ satisfies Assumption 1 in particular whenever the transfers $\beta(e^*(h))$ are a weakly increasing and weakly concave function of $e^*(h)$. Besides, as will be clear shortly, the absence of an upper bound on activity precision ($h \in \mathbb{R}_+$) 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.

Consistently with Lemma 1, precision h^* strictly increases with the weight on effort η.[22]

Proposition 1 (Equilibrium characterization). *Suppose Assumption 1 holds. Then, for $M = 0$, the unique equilibrium is all agents choosing activity h^*. By contrast, for any $M > 0$, the unique separating equilibrium (under D1) is:*

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

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

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*

[20]See Lemma 4 in Appendix A.

[21]With a discrete number of activities $N \geq 2$, multiple equilibria may coexist (under D1) due to the complementarities in the agents' choices. See Appendix **??** for the full description of the case $N = 2$.

[22]In addition, if the transfers $\beta(e^*(h))$ are monotonic with the effort $e^*(h)$ (and thus equivalently with the precision h), then $h^*/(h_0+h^*) \geq \eta$ if $\beta(e^*(h))$ increases with $e^*(h)$ (e.g. if a higher expected effort commands a higher wage), resp. $h^*/(h_0+h^*) \leq \eta$ if $\beta(e^*(h))$ decreases with h (e.g. if a higher precision is costly to implement for a university and thus requires higher tuition fees). We endogenize the transfers $\beta(e^*(h))$ in Section 2.3 via a model of competition with free entry.

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

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

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

Strikingly, except in the non-generic case $h^*/(h_0 + h^*) = \chi$, any headstart inequality $M > 0$ induces the existence of separation in equilibrium. More specifically, (i) whenever headstart image concerns dominate talent image ones (distinction case), headstart inequality generates separation "upward" with the rich distancing themselves from the poor to signal their headstart/enjoy the externalities from headstart-rich peers, whereas (ii) whenever talent image concerns dominate headstart image ones (displacement case), headstart inequality generates separation "downward" 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 headstart for the sake of their headstart image itself, whereas the displacement effect stems from the agents' desire to signal a low headstart for the sake of their talent image. The higher the headstart inequality M, the further away the agents separate.

Whether distinction or displacement prevails further depends on the weight on effort image η:[23] displacement prevails for a weight on effort η above a (possibly infinite) threshold, whereas distinction prevails below that threshold. In particular, when all activities have the same transfer β, displacement prevails if and only if $\eta > \chi$, while distinction prevails if and only if $\eta < \chi$. In other words, a stronger emphasis on effort as a component of "merit" fosters displacement.

Lastly, if χ lied outside the support of $h/(h_0+h)$ (which Assumption 1 rules out), (some) pooling would obtain in equilibrium for sufficiently high headstart inequality M as then, the chased party would be cornered at one extreme of the activity distribution, unable to escape sufficiently far away from the chasing party.[24,25]

We now turn to the agents' equilibrium payoffs.

[23] Indeed, h^* strictly increases with η, with $h^*/(h_0 + h^*) \to 1$ as $\eta \to +\infty$. Recall that η and χ are the *relative* weights of effort and headstart with respect to talent (normalized to 1).

[24] The same insights obtain with more than two headstart levels. Namely, if agents' headstart is distributed over $[0, M]$ (discretely or continuously), then (i) under distinction ($\eta < \chi$), agents with headstart 0 choose activity h^*, while agents with headstart $w \in (0, M]$ are scattered over activities with precision $h \in (h^*, \chi)$, whereas (ii) under displacement ($\eta > \chi$), agents with headstart M choose activity h^*, while agents with headstart $w \in [0, M)$ are scattered over activities with precision $h \in (\chi, h^*)$. Separation still obtains in equilibrium: whenever $\eta \neq \chi$, in any equilibrium in pure strategies (under D1), agents with different headstart levels choose different activities.

[25] When the transfers $(\beta(e^*(h)))_h$ are monotonic with $e^*(h)$, $h^* \equiv \arg\max U(h)$ satisfies $h^*/(h_0+h^*) \leq \eta$ if $\beta(e^*(h))$ decreases with $e^*(h)$, resp. $h^*/(h_0 + h^*) \geq \eta$ if $\beta(e^*(h))$ increases with $e^*(h)$. When $\eta \leq 0$ (harmful effort) and the transfers $\beta(e^*(h))$ do not sufficiently increase with expected effort $e^*(h)$, $h^* = 0$ and thus for any $\chi > 0$, distinction prevails.

Corollary 1 (Equilibrium payoffs). *Suppose Assumption 1 holds. Then, in the unique separating equilibrium (under D1),*

(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 expected payoff and the poor's increases with headstart inequality M, with the weight on effort image, η, with the weight on headstart 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 headstart χ may stem from more productive softskills, or as social capital and connections become more valuable. Then, with headstart 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 headstart χ directly disfavors them as they are poor.

Remark: Does meritocracy amplify inequality? The answer hinges on the value attributed to headstarts by other forms of social organization. As an illustration, let us consider "spot markets for performance" as the alternative to meritocracy.[26] With such spot markets for performance, in equilibrium, the difference between the rich's expected payoff and the poor's is equal to $\mu\chi M$. By contrast, with meritocracy, in a separating equilibrium, the difference is equal to $\mu[h_P/(h_0+h_P)]M > \mu\chi M$ if displacement prevails, and to $\mu[h_R/(h_0+h_R)]M < \mu\chi M$ if distinction prevails. As a consequence, with respect to spot markets for performance, meritocracy heightens inequality (and reduces effort) if displacement prevails, but mitigates inequality (and raises effort) if distinction does.

[26]Formally, let us consider the environment of Proposition 1, with a slightly modified notation: assume that "market performance" is given by

$$y^m = \theta + \eta e + \chi w + \varepsilon,$$

and that an alternative form of social organization ("spot markets for performance") rewards actual spot performance μy, whereas meritocracy rewards $\mu\big(\hat{\theta} + \eta\hat{e} + \chi\hat{w}\big)$, where $\hat{\theta}, \hat{e}, \hat{w}$ denote the audience's beliefs, i.e. meritocracy rewards the audience's expectation of the agent's performance minus the noise/luck ε. With markets-for-performance, in equilibrium, all agents exert effort e^m such that $g'(e^m) = \mu\eta$, and are indifferent over precision levels (by linearity).

2.3 Endogenous transfers (competitive equilibria with free entry)

While we focused so far on exogenous fixed transfers $(\beta_k)_k$, we endogenize the transfers (β_k) as stemming from competitive equilibria 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 (see below for purely individual tasks), 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.[27] 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 $pM > 0$, $h^*_R < h^*_a < h^*_P$.

We say that an equilibrium (for our equilibrium concept) is "in pure strategies" if all agents of a same headstart level choose the same activity.

Proposition 2 (**Endogenous rewards: Competitive equilibrium with free entry**). *For a given total headstart pM in the economy, absent headstart inequality (i.e. redistributing the total headstart pM equally across agents), the unique equilibrium is all agents choosing activity $h^*_a \in (h^*_R, h^*_P)$.*

By contrast, with headstart inequality ($w \in \{0, M\}$, $M > 0$),[28] 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

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

[28] See Appendix D for detailed existence conditions.

(under D1) exists only if

(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 may not exist for parameters regions with strictly positive measure.

3 Policy

We begin in Section 3.1 with an investigation of the optimal activity design in terms of current transfers β_k and precisions h_k, leaving future payoffs (wages or images) otherwise unchanged, i.e. by considering the optimal intervention of a principal able to ban any activity it wants (from a sufficiently large initial set) and design activity-specific current 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, we investigate the optimal taxation of (future) income and the activity-blind 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" unconditional on (then past) activity choices. Hence, Sections 3.1 and 3.2 consider direct interventions on the activities available to the agents, while Section 3.3 considers indirect interventions.

We introduce a general policy objective, which we specialize 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.[29] Namely, let the

[29]We comment in Section 3.4 on an alternative objective – *image fairness*.

principal's objective be

$$\max \left(\mathbb{E}_w\left[a_{k(w)}(e^*_{k(w)}+w)-\beta_{k(w)}\right]+\left[q_R p W_M + q_P(1-p)W_0\right]\right) \qquad (W_{q_R,q_P})$$

with a the marginal value to the principal of agents' performance, $q_R, q_P \in [0,q]$ the respective weights on the rich's and the poor's welfare, with $q \in (0,1)$,[30] and W_M and W_0 respectively a rich agent's and a poor agent's expected welfare,[31] and where $k(w)$ denotes the activity choice of an agent with headstart 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.[32]

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 $q_R = q_P = 0$, $\beta \lesseqgtr 0$ be the agents' wage or entry fee, and a the (marginal) profit from organization members' performance (which may be higher or lower than the value of the marginal long-term productivity of effort $\mu\eta$).

3.1 Optimal activity characteristics: "Second-best" meritocracy

What does the "second-best meritocracy", i.e. the optimal meritocracy for a given initial headstart inequality M, look like? We follow a mechanism design approach, assuming that the principal can design the characteristics (namely precision and fixed transfer) of as many activities as it wants.[33]

We focus on the interpretation of "headstart" 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 β subject to their participation constraint).

The principal chooses activities' precision $(h_k)_k \in \mathbb{R}^2_+$ and transfers $(\beta_k)_k \in \mathbb{R}^2$ to maxi-

[30] 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. $a = \tilde{a}/(1+\lambda)$).

[31] Namely, with our main specification of absolute image concerns,

$$W_M = \mathbb{E}\left[\beta_{k(M)} - g(e^*_{k(M)}) + \frac{\mu h_{k(M)}}{h_0 + h_{k(M)}}\big(M - \mathbb{E}[w|k(M)]\big) + \mu\eta e^*_{k(M)} + \mu\chi\mathbb{E}[w|k(M)]\right],$$

$$W_0 = \mathbb{E}\left[\beta_{k(0)} - g(e^*_{k(0)}) - \frac{\mu h_{k(0)}}{h_0 + h_{k(0)}}\mathbb{E}[w|k(0)] + \mu\eta e^*_{k(0)} + \mu\chi\mathbb{E}[w|k(0)]\right],$$

[32] Whenever the principal faces a strictly positive cost of funds, q_R, q_P are strictly below 1.

[33] Consistently with our stance on unlimited activity "capacity", we restrict our attention to deterministic mechanisms.

mize objective W_{q_R,q_P}, subject to the agents' incentive and participation constraints.[34] For simplicity, we assume that the cost of effort $g(\cdot)$ is quadratic: $g : e \mapsto g(e) = e^2/2$, that $a + \mu\eta < \mu$ (so that optimal precisions are interior). We focus on the implementation of separating equilibria.[35] that whenever several equilibra coexist, the principal can select the separating one.

The agents' outside option is another activity, beyond the principal's control, with precision $h_{out} \in \mathbb{R}_+$ and fixed transfer $\beta_{out} \in \mathbb{R}$. 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 to participate.[36]

Formally, the principal thus solves

$$\max_{\big((k(w),\ \beta_{k(w)},\ h_{k(w)})_{w\in\{0,M\}}\big)} \mathbb{E}\big[ae^*_{k(w)} - \beta_{k(w)}\big] + q_R p W_M + q_P(1-p)W_0$$

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

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

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

$$\begin{aligned}&\beta_{k(w)} + \mu\bigg(\frac{h_{k(w)}}{h_0 + h_{k(w)}}(w - \mathbb{E}[w|k(w)]) + \eta e^*_{k(w)} + \chi\mathbb{E}[w|k(w)]\bigg) - g(e^*_{k(w)})\\ &\quad\geq\ \beta_{k(w')} + \mu\bigg(\frac{h_{k(w')}}{h_0 + h_{k(w')}}(w - \mathbb{E}[w|k(w')]) + \eta e^*_{k(w')} + \chi\mathbb{E}[w|k(w')]\bigg) - g(e^*_{k(w')}).\end{aligned}$$

We refer to the *first-best precision level* h^{FB} as the one that maximizes the principal's objective absent headstart 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 headstart and on the audience's beliefs about the headstart of agents choosing the outside option. We say that incentives are *aligned* if the poor's participation constraint is binding but the rich's is not, and *countervailing* if the rich's is and the poor's is not. By the sorting condition, incentives are aligned whenever the activities' precision are higher than h_{out}, and countervailing if they are lower than h_{out}.[37]

[34] We assume that the principal can pick any positive precision. In practice, it may be much easier to add some noise than to remove some, i.e. to lower precision than to increase it. Arguably, the highest feasible precision in any given activity is finite (the lowest feasible variance is bounded away from zero). Such a constraint would yield corner solutions for the second-best activity characteristics.

[35] We conjecture that this focus is without loss of generality.

[36] I.e., such that $U_{out} + \mu\chi\mathbb{E}[w|\text{ out }]$, with $U_{out} \equiv \beta_{out} + \mu\eta e^*(h_{out}) - g(e^*(h_{out}))$, is sufficiently low.

[37] Incentive compatibility requires that the rich's activity precision be (weakly) higher) than the poor's. As a consequence, incentives are aligned whenever the poor's activitiy precision is higher than h_{out}, and countervailing if the rich's is lower than h_{out}.

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

Unprecise 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 her academic ability, and thus $h_{out} \ll h^{FB}$. Leaving the educational system is thus believed to come from "poor" agents.

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 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 non-monotonic with M for $q_R < 1$. 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 downward to reduce the rich's rent. The higher the weight on the rich's welfare, the lower the distortion.

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 $\beta_R^{SB}, \beta_P^{SB}$.

$$\begin{cases} \beta_R^{SB} = g(e^*(h_R^{SB})) - \mu\eta e^*(h_R^{SB}) + \mu\left(\dfrac{h_P^{SB}}{h_0 + h_P^{SB}} - \chi\right)M + U_{out}, \\ \beta_P^{SB} = g(e^*(h_P^{SB})) - \mu\eta e^*(h_P^{SB}) + U_{out}. \end{cases} \tag{3}$$

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 and demanding foreign universities. Leaving the national educational system for one of these foreign universities is thus believed to come from "rich" agents.

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

$$\begin{cases} \dfrac{\mu h_R^{SB}}{h_0 + h_R^{SB}} = \min\left(a + \mu\eta + (1 - q_P)\dfrac{1-p}{p}M,\ \dfrac{\mu h_{out}}{h_0 + h_{out}}\right), \\ \dfrac{\mu h_P^{SB}}{h_0 + h_P^{SB}} = a + \mu\eta. \end{cases} \tag{4}$$

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 non-monotonic with M for $q_P < 1$. The poor now exert the first-best effort level, whereas the rich exert a strictly higher effort as their activity's precision is distorted upward to reduce the poor's rent. The higher the weight on the poor's welfare, the lower the distortion.

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

$$\begin{cases} \beta_R^{SB} = g(e^*(h_R^{SB})) - \mu\eta e^*(h_R^{SB}) + U_{out}, \\ \beta_P^{SB} = g(e^*(h_P^{SB})) - \mu\eta e^*(h_P^{SB}) - \mu\Big(\dfrac{h_R^{SB}}{h_0 + h_R^{SB}} - \chi\Big)M + U_{out}. \end{cases} \tag{5}$$

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} (which depends on the weight on effort, η, and the externalities from the agents' performance, a), but also on headstart inequality M and on the weight on headstart image χ.[38]

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

Proposition 3 (**Optimal activity design**). *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 headstart inequality M, the principal chooses aligned incentives if $h_{out} < h^{FB}$, and countervailing incentives if $h_{out} > h^{FB}$. By contrast, for sufficiently large headstart 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.

[38] See Appendix E for details.

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 – whether universities or academic fields –, or stem from the outright bans of allegedly "intrinsically precise" activities – e.g. removing specific fields from school curricula (say, Latin or Ancient Greek) –, 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 decision.[39]

For simplicity, we focus on the case of a continuum of activities indexed by their precision h, and assume that Assumption 1 holds.[40] We assume that imposing a cap $\overline{h}$ on activity precision amounts to banning all activities with precision $h > \overline{h}$.

Proposition 4 (**Precision caps: Equilibrium characterization**). *Suppose that Assumption 1 holds, and 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 (under D1): the rich choose activity $\overline{h}$ and the poor choose activity $h_P(\overline{h})$ such that*

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

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 still forces the rich into an activity with lower precision, it only further displaces the poor towards an activity with an even lower precision.

Analogous insights would hold with a precision floor, which destroys all separating equilibria under displacement, but lets a separating equilibrium survive with distinction. In a sense, image concerns and headstart 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 – whether the principal's "hammer" is an activity ban or, as we will study next, income taxes.

[39] We describe in Appendix **??** targetted activity bans. The two interventions do not require the same information from the principal. Indeed, (targeted) activity bans only require a local knowledge of activities – e.g., if the principal wants to ban the rich's activity (and any closely similar activities), it needs to observe only that activity (and its close neighborhood). By contrast, enforcing a cap on activity precision requires that the principal observes for any activity, whether that activity's precision is above or below the cap.

[40] Analogous insights hold in the discrete case.

Corollary 2 (Precision caps: Separating equilibrium payoffs under displacement). *Suppose the principal sets a precision cap $\overline{h}$ in-between the rich's and the poor's laissez-faire activity precisions, $\overline{h} < h_R$). 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 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 Income taxation (and optimal intensity of image concerns)

We briefly study income taxation and emphasize a striking property of the optimal income tax in our environment (see remark below for an application to education and university funding). We leave a detailed study of optimal taxation in our image-concerns environment for future work.

For the sake of brevity, we make several simplifying assumptions. We focus on the case of a continuum of activities, indexed by their precision h, and assume that Assumption 1 holds. We further assume that the transfers $\beta(h)$ are nil,[41] and that the cost of effort g is quadratic ($g : e \mapsto e^2/2$).

For simplicity again, we suppose that the agents' effort has no externalities ($a = 0$), and the principal's objective W_{q_R,q_P} is thus a weighted sum of the rich's and the poor's welfares. We focus on settings in which headstart is non-financial and constant over time.

The principal can commit. We assume that taxation is "activity-blind" and thus cannot condition taxes and transfers 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 her expected future productivity, and hence her future wage (assuming competitive labour markets). A (persistent) income tax thus applies to the agents' future income ($\psi(y_k)$).[42]

Let us as before consider separating equilibria. With a linear income tax τ, an agent's

[41] Alternatively, we could assume that they are constant across activities. Our insights still hold in the general case with non-constant transfers $\beta(h)$, but the principal then faces additional incentives to distort the income tax in order to orient the agents towards specific activities. By contrast, we simplify our environment in order to have a clear benchmark for the optimal income tax.

[42] More generally, the income tax also applies to current income (β if positive).

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 headstart w choosing an activity with precision h has an expected payoff before redistribution given by

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

yielding tax proceeds

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

to the principal.

The principal redistributes all tax proceeds to the agents. As the principal is activity-blind, such redistribution is lump-sum. Hence, an agent with headstart w choosing an activity with precision h has a post-redistribution utility equal to

$$\begin{aligned}&(1-\tau)\left(\beta + \mu\left[\frac{h}{h_0+h}(w - \mathbb{E}[w|h])\right] + \eta e^*(h,\tau) + \chi\mathbb{E}[w|h]\right]\right) - g(e^*(h,\tau))\\ &+ \tau\mathbb{E}\left[\beta + \mu\left(\frac{h}{h_0+h}(w - \mathbb{E}[w|h])\right] + \eta e^*(h,\tau) + \chi\mathbb{E}[w|h]\right)\right]\end{aligned}$$

where the expectation is taken over the agents' precision 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 3 (**Equilibrium characterization, income tax**). *Suppose Assumption 1 holds, with β constant across activities and g quadratic. Let $\tau < 1$ be the linear income tax rate. Then, for $M = 0$, the unique equilibrium is all agents choosing activity h^*. By contrast, for any $M > 0$, the unique equilibrium (under D1) 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*

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

Hence, $h_R(\tau)$ 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*

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

Hence, $h_P(\tau)$ strictly decreases with τ.

As a consequence, absent headstart inequality ($M = 0$), the principal implements the first-best effort level with the income tax $\tau^{FB} = 0$.[43] By contrast, with headstart inequality $M > 0$ and whenever $\eta \neq \chi$, the rich and the poor separate, and by Lemma 3, 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 activities' characteristics/incentives, thereby making both parties more mobile across activities. Hence, to mitigate the impact of a higher headstart inequality, should the optimal tax decrease in order to make both the rich and the poor less mobile and reduce the distortions in activity choices?

The principal solves for $q_R = q_P = q$:

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

Our main insight then follows from the equilibrium characterization of Lemma 3.

Proposition 5 (Income taxation). *Suppose the principal puts equal weights on the rich's and the poor's welfares ($q_R = q_P = q$), and assume that its objective is concave with respect to τ. Then, absent headstart inequality ($M = 0$), the optimal income tax is nil, whereas with headstart inequality ($M > 0$), under the assumptions of Lemma 3,*

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

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

The sign and monotonicity of the optimal income tax with respect to headstart inequality thus depend on whether distinction or displacement prevails. With displacement (resp. distinction), headstart 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).

[43] 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 illustrative simplicity.

The higher the headstart 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 quality the education they can deliver. Hence, in a 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 headstart inequality, and on whether displacement or distinction prevails.[44]

Remark: Optimal intensity of image concerns (μ). Departing from 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 headstart inequality, when distinction prevails, optimal image concerns are *less intense* (lower μ) than absent headstart inequality, while when displacement prevails, they are *more intense* (higher μ) – in both cases, the more so the higher the headstart inequality. In terms of publicity of agents' merit, the higher the headstart inequality, the lower the optimal publicity if distinction prevails, but the higher the optimal publicity if displacement does.

[44]We abstract away from international competition to attract talent, which may generate/reinforce the negative consequences of a higher tax.

3.4 "Merit" and image "fairness"

Let us consider an alternative policy objective: agents having "fair" images of their individual talent, i.e. close to their actual types:[45]

$$\min\ \mathbb{E}\big[(\theta-\hat{\theta})^2\big]$$

This objective is equivalent to $\min\ \mathbb{E}\big[(\theta+\eta e-\hat{\theta}-\eta\hat{e})^2\big]$, as activity choice is public and optimal effort in a given activity does not depend on an agent's headstart. "Image fairness" could also be motivated on utilitarian grounds by invoking assortative matching considerations.

Images are closer to their true values the more the precise the agent's activity, and the closer the audience's expectation of the agent's headstart to her actual headstart. As a consequence, both the displacement and the distinction effects generate two opposite impacts:

(i) Separation screens the agents' headstart, and thus reduces the image distortions across headstart categories (rich/poor).

(ii) The displacement effect induces the poor to choose less precise activities, and thus increases the image distortions within the poor. Similarly, the distinction effect induces the rich to choose more precise activities, and thus reduces the image distortions within the rich.

Hence, whether the principal opposes or encourages displacement or distinction depends on activities' precision and on the welfare weights on the rich's and the poor's image distortions.

4 Extensions

We focus in this Section on two important extensions: dynamics and relative image concern.

4.1 Dynamics

As in our (one-period) framework, separation in activity choices perfectly reveals the agents' headstarts, one may wonder whether separation persists over time – or can 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 –

[45] A more general formulation, allowing for different weights on the rich's and the poor's image fairness, is:

$$\min\ \left(q\mathbb{E}\big[(\theta-\hat{\theta})^2|w=M\big]+(1-q)\mathbb{E}\big[(\theta-\hat{\theta})^2|w=0\big]\right)$$

e.g. high school and college, or undergrad and grad school. For simplicity, we assume that the agents' headstart remains constant over the two periods – e.g. because periods are short and headstarts are non-financial. We refer to *separating equilibria* as equilibria in which the agents' headstart is perfectly revealed (either in the first or the second period).

We show that in this two-period environment, separation can obtain in the first period, thereby perfectly revealing each agent's headstart 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 to 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. They receive a period-1 transfer β_1 (positive or negative) associated to activity h_1, and incur their 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, receiving a period-2 transfer β_2 and 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 her expected talent, sum of expected (past) efforts and expected headstart given I – e.g. in a competitive recruitment environment, an agent's image concern is (the discounted sum of) her future wage(s).

An agent with publicly observable activity choices and outcomes $\{h_1, y_1, h_2, y_2\} \equiv I$ thus maximizes[46]

$$\mathbb{E}\Big[\beta_1 - g(e_1) + \delta[\beta_2 - g(e_2)] + \delta\mu\Big(\mathbb{E}[\theta|I] + \eta(\mathbb{E}[e_1|I] + \mathbb{E}[e_2|I]) + \chi\mathbb{E}[w|I]\Big)\Big]$$

In any 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 separating equilibrium is thus given by

$$g'(e_2^*) = \frac{\mu h_2}{h_0 + h_1 + h_2}. \tag{6}$$

[46]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.

For simplicity, we assume that $\eta \in (0, 1/2)$, which will yield interior solutions.[47]

Agents may be tempted by "double deviations", combining a deviation in their period-1 effort level e_1 with one in their period-2 activity choice h_2.[48] For simplicity, we assume passive beliefs. Namely, we assume that the audience does not update its beliefs $\hat{e}_1$ regarding an individual's period-1 effort based on the individual's period-1 performance (y_1), period-2 precision choice (h_2) nor period-2 performance (y_2), i.e. that the audience only uses the agent's period-1 precision choice (h_1) to form its beliefs about e_1.[49] In addition, in order to minimize the role of self-fulfilling beliefs, we assume that when an agent's optimal activity choice h_t, resp. effort e_t in period $t \in \{1, 2\}$ does not depend directly on her own headstart (but only possibly indirectly via the audience's belief), the audience's belief regarding the agent's headstart does not depend on the agent's actual activity choice h_t, resp. period-t outcome y_t.

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 $e_1^*(h_1, h_2^*)$ is given by

$$g'(e_1^*) = \frac{\delta\mu h_1}{h_0 + h_1 + h_2^*(h_1)}. \tag{7}$$

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 6 (Equilibrium characterization, two-period model). *Suppose agents choose among a continuum of activities indexed by their precision $h \in \mathbb{R}_+$, with period-t transfer β_t, $t = 1, 2$, constant across activities, and a quadratic cost of effort $g(e) = e^2/2$. Suppose the audience has passive beliefs. Then, for $M = 0$, the unique equilibrium in pure strategies (under our equilibrium refinement) is all agents choosing activity h_1^* in period 1 and activity $h_2^*(h_1^*)$ in period 2. By contrast, for any $M > 0$, the unique separating equilibrium*

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

[48] Indeed, the information available to the audience at the end of period 1 is thus $\{h_1, y_1\}$, while the agent knows $\{h_1, y_1, e_1\}$ (perfect recalling her effort e_1).

[49] A rationale is that the audience forms its beliefs using only the information available to the agent at the time she took her decision.

in pure strategies (under our equilibrium refinement) is:

(i) ***(Distinction)*** *If $2\eta < \chi$, the separating equilibrium in which in period 1 the poor choose activity h_1^* while the rich choose 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 $2\eta > \chi$, it is the separating equilibrium in which 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 6 shows that with initial headstart inequality, there exist equilibria with immediate and persistent separation.[50] Separation measured by the distance between activity precisions decreases over time.

4.2 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 (pure) prestige concerns, yet are also consistent with a career-concerns interpretation, in which an agent's image is her future wage but the utility the agent ultimately derives from her future wage depends on how the latter compares to the others, or in which an agent's chances of being promoted depend on her relative "qualities" with respect to her rivals'.[51]

In the spirit of Merton (1957), we distinguish two *reference groups* for each agent: the whole society, and the agent's activity peers.[52] The agents' weights on each reference group may stem from society's division along (or mobility across) activity lines. we emphasize that

[50]Separation is here measured in terms of activity choice (i.e. activity precision). However, in period 2, both the rich and the poor exert the same effort level ($(g')^{-1}(\mu\eta)$).

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

[52]We assume that even with relative image concerns, the agent's utility from her direct reward does not depend on how it compares to other agents' direct rewards. Our insights are robust to such comparisons. A rationale for our assumption is that the instant utility that the agent derives from her current transfer β accrues before she engages in any comparisons with other agents, whereas her image/career concerns involve future benefits (prestige/wage) and she cannot avoid engaging in comparisons over time.

milder across-activity image concerns – e.g. due to a more divided, less mobile society – tend to foster displacement.

Formally, suppose that each agent cares about a weighted sum of her *local*, i.e. within-activity image $\psi(y_k) - \mathbb{E}[\psi(y_k)]$, and her *global*, i.e. across-activity image $\psi(y_k) - \mathbb{E}[\psi(y)]$ (where the expectation is taken over all activities, given the anticipated characteristics and effort levels of agents taking each activity):

$$(1-\zeta)\Big(\psi(y_k) - \mathbb{E}[\psi(y_k)]\Big) + \zeta\Big(\psi(y_k) - \mathbb{E}[\psi(y)]\Big) \tag{8}$$

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-activities may depend in particular on the extent to which activities, careers or society more generally are clustered. For instance, as a side-product of exerting effort in a given activity, agents may learn an activity-specific knowledge – either technical or relative to a profession's/a firm's culture/social norms. If such knowledge is valuable only in that given activity and worthless in others, image concerns are purely within-activity as career concerns are exclusively within-activity (agents compete only with their activity fellows). By contrast, if there is no such activity-specific knowledge, image and career concerns are across-activity as agents compete with all other agents across activities.[53] Similarly, if society is clustered along activity lines, so that agents in different activities have few interactions, if any – living and working in different neighbourhoods, having different lifestyles, etc. –, images may be mainly within-activity as agents put a higher weight on their comparisons with respect to their activity-peers – who are in this case also likely to be their neighbours, etc.

Proposition 7 (**Relative image concerns and reference groups: Results under *lais-***

[53] Mobility across activities may be asymmetric. For instance, in the course of one's career, it may be relatively easy to move from activity 1 to activity 2, e.g. from academia to the industry, but much more difficult to move in the other direction. Such asymmetric mobility would generate "path-specific" image concerns, i.e. parameters ζ_{kl} for the concerns of an agent currently in activity k with respect to those in activity l – so that in our example, we may have $\zeta_{12} > 0 = \zeta_{21}$. The activity-$k$ agent's overall career concerns could write as

$$\Big(1 - \sum_l \zeta_{kl}\Big)\Big(\psi(y_k) - \mathbb{E}[\psi(y)|\eta_{k'k} > 0]\Big) + \sum_l \zeta_{kl}\Big(\psi(y_k) - \mathbb{E}[\psi(y)|\eta_{k'l} > 0]\Big).$$

where images take into account the comparison with all potential future activity-l agents, i.e. currently in an activity k' such that $\zeta_{k'l} > 0$ – e.g. agents in activity 1 face future potential competition only from agents currently in activity 2, while agents in activity 2 face future potential competition from all agents.

Analogous insights would obtain. Following on our example, activity 2 would become less attractive, as the effort exerted in this activity would have a lower value to recruiters in other activities.

***sez-faire*).** *All formal results in Section 2, namely Lemmas 1 & 2, Proposition 1, Corollary 1 and Proposition 2 hold with general relative image concerns (as defined by (8)), 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. Indeed, as a society becomes more divided along activity lines – e.g. as education or the organization of work becomes more specialized and students or workers in different activities (fields, functions, jobs) interact less –, across-activities comparisons may matter less, i.e. 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-activities comparison may be heightened and ζ may increase.

For simplicity, suppose transfers β are constant across activities – as a consequence, with Assumption 1, 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 $\eta < \chi$).

Corollary 3 (Distinction and displacement in divided vs mobile societies). *Suppose Assumption 1 holds, and that transfers* β *are constant across activities. Then, the more mobile the society (the higher* ζ*), the larger the magnitude of distinction and the smaller the magnitude of displacement.*[54] *Conversely, the more divided the society (the lower* ζ*), the smaller the magnitude of distinction and the larger the magnitude of displacement.*

The regions of displacement and distinction may change in general with ζ. In particular, if agents can only choose their activities within a range of precisions bounded below, then whenever ζ is sufficiently low, distinction disappears: in a (sufficiently) highly divided society, only displacement prevails. For an application, consider education – e.g. with students choosing a major or a university. Then, under the above conditions, whenever teaching/knowledge is highly specialized or society strongly divided along educational lines, headstart inequality always induces displacement.

5 Related literature

Theoretical literatures. This paper builds on several theoretical literatures, too vast to be summarized here. The first founding one 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

[54] 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.

activities' information structure, allowing for complementarities between talent and effort,[55] and applying their analysis to a multitasking environment.[56] 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, 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 we introduce an (initial) privately observable heterogeneity, which as we show, distorts the agents' equilibrium effort choices, generating displacement or distinction.[57]

The second founding 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 headstart. The complementarity between the agents' headstart and their precision choices, which generates the sorting condition, arises endogenously from the agents' effort to influence their talent image. Our agents' trade-off between talent, effort and headstart images can be compared to 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 only know one dimension of their type (headstart) and face a single audience to which they send a

[55] Relatedly, while most of the literature on career concerns assumes that agents' performances are observable throughout, 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).

[56] Cisternas (2018) introduces strategic skill acquisition, studying environemnts in which effort is a direct input both to current production and to skill acquisition, and evidences 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 headstart leads to suboptimally low effort (from the poor) when displacement prevails, but suboptimally high effort (from the rich) when distinction does.

[57] 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 their 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' headstarts 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 4.2). 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).

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 the privately observable headstarts and unfolds across activity choices. In addition, our results on the optimal intensity of image concerns are related to Ali and Bénabou (2020), who evidence 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) organization, such as Coase (1937), but also Cole, Mailath and Postlewaite (1992), Burdett and Coles (1997), and Fernandez and Gali (1999) who compare 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. In particular, as mentioned in Section 2.2, it induces lower effort and higher payoff inequality than spot markets for performance if displacement prevails, but higher effort and lower payoff inequality if distinction does. The matching efficiency of meritocracy – the accuracy of the audience's beliefs on the agents' type – 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 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/failures (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 addition to objective headstarts 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, i.e. more "elite" 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.[58] In our model, the agents' differentiated choices obtain even in the absence of any such narrative or class-specific aspirations.

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 one of middle-class applicants and decreased the one of lower-SES applicants. From our model perspective, this may suggest that (some degree of) distinction prevailed between the middle class and lower-SES applicants (the higher precision allowing the middle class to further separate further), and that (some degree of) displacement prevailed between the middle class and higher-SES applicants (the higher precision making the higher-SES applicants willing to stay in that activity).[59]

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 distinction in our model.[60] Macchi (2023) provides evidence on credit-worthiness signalling strategies

[58] 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 their child's prospective application to a given university may reveal to the child how much parental support (material and immaterial) she can expect were she to attend that university. Such parental support constitutes a headstart (privately observable and affecting the student's performance). While these two explanations point to the same outcome, they call for different remedies.

[59] Our model could be extended to deliver such differentiated outcomes. Let us sketch such an extension. Suppose headstarts 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 headstart dimension χ_1, χ_2. Suppose there are three "classes" in the population: the upper class with headstart (M, M), the middle class with headstart $(0, M)$ and the lower class with headstart $(0, 0)$. (Suppose, for simplicity, that agents face a continuum of activities and that transfers β do not depend on precision.) 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. Separation remains the unique equilibrium outcome. By contrast, if $\chi_1 < \eta < \chi_2$, pooling equilibria can exist.

[60] Interestingly, Bursztyn et al (2018) evidence 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

in Uganda, showing that obesity facilitates credit access. Interpreting these strategies as aimed at signalling not only wealth (privately observed "headstart" in our model) but also reliability (unobserved "talent"), the outcome may correspond to displacement.[61]

6 Alleys for future research

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

Headstarts and occupational change. Could the simple model we introduced in this paper be extended to explain occupational change, and in particular, 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 career occupations (see e.g. Autor 2019)? Such an extension may require introducing multidimensional headstarts (as sketched in the previous Section), then asking which changes in the weights on talent or different dimensions of headstarts could explain the patterns observed in the data.

Monetary headstarts and optimal taxation. For expositional simplicity, we focused in this paper on non-monetary headstarts – human or social capital. However, headstarts may have a monetary component – even if indirect, e.g. ability to pay for private tutoring or summer camps, or for more comfortable or healthier living conditions, etc. How would optimal policies change given this monetary component? In particular, taxes (or subsidies) may allow to redistribute part of the headstarts across individuals and across generations.[62]

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 suppply equilibrium. Yet, as pointed by moralists and philosophers (see e.g. Sandel 2020), markets and moral narratives may put different weights on each component of "merit" – innate talent, effort and headstart. Could the model shed some light on the (joint) relations between different production technologies, modes of organization and moral narratives?[63]

---

prevails (see Section 3.3).

[61] Importantly, Macchi (2023) shows indeed that while obesity is correlated with wealth, it is not interpreted as a signal of beauty or health, which may suggest that if they could, individuals may otherwise prefer other signals of credit-worthiness.

[62] Such a line of research would be related to the literature on dynamic public finance and optimal estate taxation, such as Farhi and Werning (2010).

[63] Such a line of research would lie in the wake of Weber's (1905) seminal work on "the protestant ethic and the spirit of capitalism".

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.*

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.

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

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.

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

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.

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.

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.

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.

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

Langtry, A. (2022). "Keeping Up with the "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.

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.

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

We explicit our equilibrium refinement, which is in the spirit of D1. We write the definition for separating equilibria. Let the audience's beliefs be described by the vector $p' \equiv (p'_k)_k$ where $p'_k \in \{0,1\}$ is the probability attributed to an agent choosing activity k being rich. Fix an equilibrium, and consider an out-of-equilibrium activity choice k. Denote by $u^*(w)$ the expected equilibrium payoff of an agent with headstart w, and by $u(w, p'_k, k)$ the expected payoff of an agent with headstart w when choosing activity k given the audience's beliefs p'. Lastly, for any headstart level $w \in \{0, M\}$, let $D_k(w) \equiv \{p'_k \in \{0,1\} \mid u^*(w) < u(w, p'_k, k)\}$ and $D^0_k(w) \equiv \{p'_k \in \{0,1\} \mid u^*(w) = u(w, p'_k, k)\}$.

Definition. ***(Equilibrium concept, separating equilibria)*** *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 k and for a headstart level w, there exists another headstart level w' such that $D_k(w) \cup D^0_k(w) \subseteq D_k(w')$, then the audience's equilibrium belief should put zero probability on an agent who chooses activity k having headstart w.*

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

Lemma 4. ***(Sorting and off-path beliefs)*** *Suppose there exists a separating equilibrium 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 $p(h)$ denote the equilibrium 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,

$$\begin{cases} U(h_P) - \mu\Big(\dfrac{h_P}{h_0+h_P} - \chi\Big)p(h_P)M \ \geq\ U(h_R) - \mu\Big(\dfrac{h_R}{h_0+h_R} - \chi\Big)p(h_R)M, \\ U(h_R) + \dfrac{\mu h_R}{h_0+h_R}M - \mu\Big(\dfrac{h_R}{h_0+h_R} - \chi\Big)p(h_R)M \\ \qquad\qquad \geq\ U(h_P) + \dfrac{\mu h_P}{h_0+h_P}M - \mu\Big(\dfrac{h_P}{h_0+h_P} - \chi\Big)p(h_P)M, \end{cases} \tag{9}$$

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 (9) 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, (9) implies that the equilibrium belief $p(h)$ for any off-path activity $h > h_R$ is equal to 1. □

B Proofs of Lemmas 1 and 2

B.1 Proof of Lemma 1

Agents choose their activity by solving $\max\limits_k U_k$ where

$$U_k = \mu\eta e_k^* + \beta_k - g(e_k^*)$$

and e_k^* is given by (1). Let $H_k \equiv h_k/(h_0 + h_k)$. By differentiation and using (1),

$$\frac{\partial U_k}{\partial H_k} = \frac{\mu^2(\eta - H_k)}{g'' \circ (g')^{-1}(\mu H_k)} = \frac{\mu^2(\eta - H_k)}{g''(e_k^*)}.$$

The results follow.

B.2 Proof of Lemma 2

A rich agent's (expected) payoff in activity k is given by

$$V_k(M) \equiv U_k + \frac{\mu h_k}{h_0 + h_k} M - \mu\Big(\frac{h_k}{h_0 + h_k} - \chi\Big)\mathbb{E}[w|k],$$

while a poor agent's (expected) payoff in activity k is given by

$$\begin{aligned} V_k(0) &\equiv U_k - \mu\Big(\frac{h_k}{h_0 + h_k} - \chi\Big)\mathbb{E}[w|k] \\ &= V_k(M) - \frac{\mu h_k}{h_0 + h_k} M. \end{aligned}$$

Hence, for any $M > 0$,

$$\frac{\partial}{\partial h_k} V_k(M) > \frac{\partial}{\partial h_k} V_k(0).$$

C Proof of Proposition 1 and Corollary 1

Suppose Assumption 1 holds. Hence, $U(h^*) > U(h)$ for any $h \neq h^*$, and therefore, absent headstart 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 4, 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[64]

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

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 4, $h_P < h_R$. Our preliminary remark together with Assumption 1 yield that

$$h_P \leq h^* \qquad \text{and} \qquad h_R \geq h^*. \tag{10}$$

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:

$$\mu\Big(\frac{h_P}{h_0 + h_P} - \chi\Big)M \leq U(h_R) - U(h_P) \leq \mu\Big(\frac{h_R}{h_0 + h_R} - \chi\Big)M. \tag{11}$$

With our equilibrium concept, off-path deviations to an activity $h \in (h_P, h_R)$ are attributed 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 (11) implies that $h' \in [h_P, h_R]$, and the necessary and sufficient existence

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

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):[65]

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

$$\mu\Big(\frac{h_R}{h_0 + h_R} - \chi\Big)M = U(h_R) - U(h^*). \tag{12}$$

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

$$\mu\Big(\frac{h_P}{h_0 + h_P} - \chi\Big)M = U(h^*) - U(h_P). \tag{13}$$

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.

Monotonicity of h_P and h_R. The monotonicity of h_P and h_R with respect to M obtain with the implicit function theorem by differentiating (13) and (12), as by Assumption 1, the function U strictly increases with $h \in (0, h^*)$ and strictly decreases with $h \in (h^*, +\infty)$.

D Proof of Proposition 2

Let us prove the following result, from which Proposition 2 immediately follows.

Proposition 8. *For a given total headstart in the economy pM, absent headstart inequality (i.e. redistributing the total headstart pM equally across agents), the unique equilibrium is all agents choosing activity $h_a^* \in (h_R^*, h_P^*)$. By contrast, with headstart inequality ($w \in \{0, M\}$, $M > 0$), a separating equilibrium in pure strategies (under D1) exists if and only if either*

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

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

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

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 $p : h \longmapsto p(h)$ be the belief function such that an agent in activity with precision $h \geq 0$ is believed to be rich with probability $p(h) \in [0, 1]$. As noted in the text, free entry and wage competition yield that firms choosing precision h (if any) offer a wage $\pi\big(e^*(h)+p(h)M\big)$. 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 $U(h) \equiv \mu\eta e^*(h) - g(e^*(h))$.

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 4, 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[66]

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

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

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 4 (see Appendix A), $h_R < h_P$. Our preliminary remark yields that

$$h_P \leq h_P^* \qquad \text{and} \qquad h_R \geq h_R^*. \tag{14}$$

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:

$$\mu\Big(\frac{h_P}{h_0 + h_P} - \chi\Big)M \leq U(h_R) - U(h_P) + \pi(e^*(h_R) + M) - \pi(e^*(h_P)) \;\leq\; \mu\Big(\frac{h_R}{h_0 + h_R} - \chi\Big)M. \tag{15}$$

With our equilibrium concept, off-path deviations to an activity $h \in (h_P, h_R)$ are attributed to poor agents with probability 1 if for any belief $p \in \{0, 1\}$,

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

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 (15) 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").[67]

E Proof of Proposition 3

We begin with a remark regarding the audience's beliefs $\mathbb{E}[w|\text{out}]$. By Lemma 2, with our equilibrium concept, $\mathbb{E}[w|\text{out}] = M$ if $h_{out} > h_{k(M)}$ and $\mathbb{E}[w|\text{out}] = 0$ if $h_{out} < h_{k(0)}$. In words, a deviation to the outside option is attributed to a rich agent if the outside option is

[67]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 $h_P^* > h_R^*$,

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

more precise than the rich's activity $k(M)$, and to a poor agent if the outside option is less precise than the poor's activity $k(0)$.

To alleviate the notation in the rest of this Section (only), we index by 1 the rich's activity and by 2 the poor's.

Aligned incentives. We define aligned incentives as the case in which deviations to the outside option are attributed to poor agents: $\mathbb{E}[w|\text{out}] = 0$. As discussed in our preliminary remark, aligned incentives thus require that $h_{k(0)} \geq h_{out}$.

Hence, with aligned incentives, i.e. subject to the constraint $\mathbb{E}[w|\text{out}] = 0$, standard arguments yield that the optimal activity characteristics are given by:

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

$$\begin{cases} \dfrac{\mu h_1}{h_0 + h_1} = a + \mu\eta, \\ \dfrac{\mu h_2}{h_0 + h_2} = \max\left(\dfrac{\mu h_{out}}{h_0 + h_{out}},\ a + \mu\eta - (1 - q_R)\dfrac{p}{1-p}M\right) \end{cases}$$

for precisions, and for transfers

$$\begin{cases} \beta_1 = g(e_1^*) - \mu\eta e_1^* + \mu\left(\dfrac{h_2}{h_0 + h_2} - \chi\right)M + U_{out}, \\ \beta_2 = g(e_2^*) - \mu\eta e_2^* + U_{out}. \end{cases}$$

(ii) if $a + \mu\eta \leq \mu h_{out}/(h_0 + h_{out})$,

$$\frac{\mu h_1}{h_0 + h_1} = \frac{\mu h_2}{h_0 + h_2} = \frac{\mu h_{out}}{h_0 + h_{out}},$$

for precisions, and for transfers

$$\begin{cases} \beta_1 = \beta_{out} + \mu\left(\dfrac{h_{out}}{h_0 + h_{out}} - \chi\right)M, \\ \beta_2 = \beta_{out}. \end{cases}$$

Countervailing incentives. We define countervailing incentives as the case in which deviations to the outside option are attributed to rich agents: $\mathbb{E}[w|\text{out}] = M$. As discussed in our preliminary remark, countervailing incentives thus require that $h_{k(M)} \leq h_{out}$.

Hence, with aligned incentives, i.e. subject to the constraint $\mathbb{E}[w|\text{out}] = 0$, standard arguments yield that the optimal activity characteristics are given by:

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

$$\begin{cases} \dfrac{\mu h_1}{h_0 + h_1} = \min\left(a + \mu\eta + (1 - q_P)\dfrac{1-p}{p}M,\ \dfrac{\mu h_{out}}{h_0 + h_{out}}\right), \\ \dfrac{\mu h_2}{h_0 + h_2} = a + \mu\eta. \end{cases}$$

for precisions, and for transfers

$$\begin{cases} \beta_1 = g(e_1^*) - \mu\eta e_1^* + U_{out}, \\ \beta_2 = g(e_2^*) - \mu\eta e_2^* - \mu\Big(\dfrac{h_1}{h_0 + h_1} - \chi\Big)M + U_{out}. \end{cases}$$

(ii) if $a + \mu\eta \geq \mu h_{out}/(h_0 + h_{out})$,

$$\frac{\mu h_1}{h_0 + h_1} = \frac{\mu h_2}{h_0 + h_2} = \frac{\mu h_{out}}{h_0 + h_{out}},$$

for precisions, and for transfers

$$\begin{cases} \beta_1 = \beta_{out}, \\ \beta_2 = \beta_{out} - \mu\Big(\dfrac{h_{out}}{h_0 + h_{out}} - \chi\Big)M. \end{cases}$$

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

If $h_{out} < h^{FB}$, the principal compares the optimal activity characteristics conditional on aligned incentives, which yield[68]

$$p\big[(a + \mu\eta)e^*(h_1) - g(e^*(h_1))\big] + (1 - p)\big[(a + \mu\eta)e^*(h_2) - g(e^*(h_2))\big] - (1 - q_R)p\frac{\mu h_2}{h_0 + h_2}M + p\mu\chi M$$

where h_1 and h_2 are given by (2), with those conditional on countervailing incentives, which yield

$$(a + \mu\eta)e^*(h_{out}) - g(e^*(h_{out})) + (1 - q_P)(1 - p)\mu\Big(\frac{h_{out}}{h_0 + h_{out}} - \chi\Big)M + q_R p\mu\chi M.$$

[68]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.

In particular, for M sufficiently high that (2) yields $h_2 = h_{out}$, the principal chooses aligned incentives if and only if

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

and countervailing incentives otherwise.

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

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

with

$$p\big[(a+\mu\eta)e^*(h_1) - g(e^*(h_1))\big] + (1-p)\big[(a+\mu\eta)e^*(h_2) - g(e^*(h_2))\big]$$
$$+ (1-q_P)(1-p)\mu\Big(\frac{h_1}{h_0+h_1} - \chi\Big)M + q_R p\mu\chi M$$

where h_1 and h_2 are given by (4). In particular, for M sufficiently high that (4) yields $h_1 = h_{out}$, the principal chooses aligned incentives if and only if

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

and countervailing incentives otherwise.

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

F Proofs of Proposition 4 and Corollary 2

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

Suppose that Assumption 1 holds, and 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 5

Lemma 3 follows from Proposition 1, and its proof is thus omitted.

Suppose Assumption 1 holds, with β constant across activities and g quadratic. For any

$\tau \leq 1$, let

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

The principal thus solves:

$$\max_{\tau} \ W(\tau).$$

Suppose $W(\tau)$ is concave with respect to τ.

Differentation then yields that

$$\frac{\partial W}{\partial \tau} = p\Big(\mu\eta - (1-\tau)\frac{\mu h_R}{h_0 + h_R}\Big)\Big(-\frac{\mu h_R}{h_0 + h_R} + (1-\tau)\frac{\mu h_0}{(h_0 + h_R)^2}\frac{\partial h_R}{\partial \tau}\Big) + (1-p)\Big(\mu\eta - (1-\tau)\frac{\mu h_P}{h_0 + h_P}\Big)\Big(-\frac{\mu h_P}{h_0 + h_P} + (1-\tau)\frac{\mu h_0}{(h_0 + h_P)^2}\frac{\partial h_P}{\partial \tau}\Big) \tag{16}$$

and

$$\begin{aligned} \frac{\partial^2 W}{\partial \tau^2} = & \\ & - p\Big(\frac{\mu h_R}{h_0 + h_R} - (1-\tau)\frac{\mu h_0}{(h_0 + h_R)^2}\frac{\partial h_R}{\partial \tau}\Big)^2 \\ & - p\Big(\mu\eta - (1-\tau)\frac{\mu h_R}{h_0 + h_R}\Big)\Big[\frac{2\mu h_0}{(h_0 + h_R)^2}\frac{\partial h_R}{\partial \tau} + 2(1-\tau)\frac{\mu h_0}{(h_0 + h_R)^3}\Big(\frac{\partial h_R}{\partial \tau}\Big)^2 \\ & \qquad - (1-\tau)\frac{\mu h_0}{(h_0 + h_R)^2}\frac{\partial^2 h_R}{\partial \tau^2}\Big] \\ & - (1-p)\Big(\frac{\mu h_P}{h_0 + h_P} - (1-\tau)\frac{\mu h_0}{(h_0 + h_P)^2}\frac{\partial h_P}{\partial \tau}\Big)^2 \\ & - (1-p)\Big(\mu\eta - (1-\tau)\frac{\mu h_P}{h_0 + h_P}\Big)\Big[\frac{2\mu h_0}{(h_0 + h_P)^2}\frac{\partial h_P}{\partial \tau} + 2(1-\tau)\frac{h_0}{(h_0 + h_P)^3}\Big(\frac{\partial h_P}{\partial \tau}\Big)^2 \\ & \qquad - (1-\tau)\frac{\mu h_0}{(h_0 + h_P)^2}\frac{\partial^2 h_P}{\partial \tau^2}\Big] \end{aligned} \tag{17}$$

and

$$
\begin{aligned}
&\frac{\partial^2 W}{\partial M \partial \tau} = \qquad (18)\\
&- p(1-\tau)\frac{\mu h_0}{(h_0+h_R)^2}\frac{\partial h_R}{\partial M}\Big(-\frac{\mu h_R}{h_0+h_R}+(1-\tau)\frac{\mu h_0}{(h_0+h_R)^2}\frac{\partial h_R}{\partial \tau}\Big)\\
&- p\Big(\mu\eta-(1-\tau)\frac{\mu h_R}{h_0+h_R}\Big)\Big[\frac{\mu h_0}{(h_0+h_R)^2}\frac{\partial h_R}{\partial M}+2(1-\tau)\frac{\mu h_0}{(h_0+h_R)^3}\frac{\partial h_R}{\partial M}\frac{\partial h_R}{\partial \tau}\\
&\qquad\qquad - (1-\tau)\frac{\mu h_0}{(h_0+h_R)^2}\frac{\partial^2 h_R}{\partial M\partial \tau}\Big]\\
&- (1-p)(1-\tau)\frac{\mu h_0}{(h_0+h_P)^2}\frac{\partial h_P}{\partial M}\Big(-\frac{\mu h_P}{h_0+h_P}+(1-\tau)\frac{\mu h_0}{(h_0+h_P)^2}\frac{\partial h_P}{\partial \tau}\Big)\\
&- (1-p)\Big(\mu\eta-(1-\tau)\frac{\mu h_P}{h_0+h_P}\Big)\Big[\frac{\mu h_0}{(h_0+h_P)^2}\frac{\partial h_P}{\partial M}+2(1-\tau)\frac{\mu h_0}{(h_0+h_P)^3}\frac{\partial h_P}{\partial M}\frac{\partial h_P}{\partial \tau}\\
&\qquad\qquad - (1-\tau)\frac{\mu h_0}{(h_0+h_P)^2}\frac{\partial^2 h_P}{\partial M\partial \tau}\Big]
\end{aligned}
$$

Distinction. Suppose $\eta < \chi$. By Lemma 3, $h_P/(h_0+h_P) = \eta$ and

$$(1-\tau)\Big(\mu\eta - \frac{1}{2}\frac{\mu h_R}{h_0+h_R}\Big)\frac{\mu h_R}{h_0+h_R} = \frac{1}{2}(1-\tau)(\mu\eta)^2 + \mu\Big(\frac{h_R}{h_0+h_R}-\chi\Big)M.$$

Therefore,

$$\frac{\partial h_P}{\partial \tau} = \frac{\partial h_P}{\partial M} = 0,$$

while

$$
\begin{aligned}
&\Big[(1-\tau)\Big(\mu\eta-\frac{\mu h_R}{h_0+h_R}\Big)\frac{\mu h_0}{(h_0+h_R)^2} - \frac{\mu h_0}{(h_0+h_R)^2}M\Big]\frac{\partial h_R}{\partial \tau}\\
&\qquad = \Big(\mu\eta - \frac{1}{2}\frac{\mu h_R}{h_0+h_R}\Big)\frac{\mu h_R}{h_0+h_R} - \frac{(\mu\eta)^2}{2} = -\frac{1}{2}\Big(\frac{\mu h_R}{h_0+h_R}-\mu\eta\Big)^2,
\end{aligned}
$$

and

$$\Big[(1-\tau)\Big(\mu\eta-\frac{\mu h_R}{h_0+h_R}\Big)\frac{\mu h_0}{(h_0+h_R)^2} - \frac{\mu h_0}{(h_0+h_R)^2}M\Big]\frac{\partial h_R}{\partial M} = (1-\tau)\mu\Big(\frac{h_R}{h_0+h_R}-\chi\Big),$$

and thus

$$
\begin{aligned}
&\Big[-(1-\tau)\Big(\frac{\mu h_0}{(h_0+h_R)^2}\Big)^2 - 2(1-\tau)\Big(\mu\eta-\frac{\mu h_R}{h_0+h_R}\Big)\frac{\mu h_0}{(h_0+h_R)^3} + \frac{2\mu h_0}{(h_0+h_R)^3}M\Big]\frac{\partial h_R}{\partial M}\frac{\partial h_R}{\partial \tau}\\
&\qquad + \Big[(1-\tau)\Big(\mu\eta-\frac{\mu h_R}{h_0+h_R}\Big)\frac{\mu h_0}{(h_0+h_R)^2} - \frac{\mu h_0}{(h_0+h_R)^2}M\Big]\frac{\partial^2 h_R}{\partial M\partial \tau}\\
&= \frac{\mu h_0}{(h_0+h_R)^2}\frac{\partial h_R}{\partial \tau} + \Big(\mu\eta - \frac{\mu h_R}{h_0+h_R}\Big)\frac{\mu h_0}{(h_0+h_R)^2}\frac{\partial h_R}{\partial M}. \qquad (19)
\end{aligned}
$$

Hence, in particular, (16) writes as

$$\begin{aligned}\frac{\partial W}{\partial \tau} &= p\Big(\mu\eta - (1-\tau)\frac{\mu h_R}{h_0+h_R}\Big)\Big(-\frac{\mu h_R}{h_0+h_R} + (1-\tau)\frac{\mu h_0}{(h_0+h_R)^2}\frac{\partial h_R}{\partial \tau}\Big) - (1-p)\tau(\mu\eta)^2 \\ &= p\Big(\mu\eta - (1-\tau)\frac{\mu h_R}{h_0+h_R}\Big)\frac{\frac{(1-\tau)}{2}\Big(\Big(\frac{\mu h_R}{h_0+h_R}\Big)^2 - (\mu\eta)^2\Big) + \frac{\mu h_R}{h_0+h_R}M}{(1-\tau)\Big(\mu\eta - \frac{\mu h_R}{h_0+h_R}\Big) - M} - (1-p)\tau(\mu\eta)^2.\end{aligned} \tag{20}$$

By Lemma 3, $\mu\eta < \mu h_R/(h_0+h_R)$. Therefore, $\partial h_R/\partial\tau > 0$, and from (20), any solution τ^* to $\partial W(\tau)/\partial\tau = 0$ is necessarily strictly positive, $\tau^* > 0$, and such that $\mu\eta < (1-\tau^*)\mu h_R/(h_0+h_R)$.

In addition, the above computations imply that $h_R/(h_0+h_R)$ goes to χ as τ goes to 1. Hence, $\partial W(\tau)/\partial\tau$ goes to $-p\mu\eta\chi - (1-p)(\mu\eta)^2 < 0$ as τ goes to 1. As a consequence, any solution τ^* to $\partial W(\tau)/\partial\tau = 0$ is such that $\tau^* \in (0,1)$.[69]

[69] Equation (17) writes as

$$\begin{aligned}&\frac{\partial^2 W}{\partial\tau^2} \\ &= -(1-p)(\mu\eta)^2 - p\Big(\frac{\mu h_R}{h_0+h_R} - (1-\tau)\frac{\mu h_0}{(h_0+h_R)^2}\frac{\partial h_R}{\partial\tau}\Big)^2 \\ &\quad - p\Big(\mu\eta - (1-\tau)\frac{\mu h_R}{h_0+h_R}\Big)\Big[\frac{2\mu h_0}{(h_0+h_R)^2}\frac{\partial h_R}{\partial\tau} + 2(1-\tau)\frac{\mu h_0}{(h_0+h_R)^3}\Big(\frac{\partial h_R}{\partial\tau}\Big)^2 \\ &\qquad - (1-\tau)\frac{\mu h_0}{(h_0+h_R)^2}\frac{\partial^2 h_R}{\partial\tau^2}\Big]\end{aligned}$$

and thus in particular, for τ^* such that $\partial W(\tau^*)/\partial\tau^* = 0$,

$$\begin{aligned}&\frac{\partial^2 W}{\partial\tau^2} \\ &= -(1-p)(\mu\eta)^2 - \frac{(1-p)^2}{p}\Bigg(\frac{\tau(\mu\eta)^2}{\mu\eta - (1-\tau)\frac{\mu h_R}{h_0+h_R}}\Bigg)^2 \\ &\quad - p\Big(\mu\eta - (1-\tau)\frac{\mu h_R}{h_0+h_R}\Big)\Big[\frac{2\mu h_0}{(h_0+h_R)^2}\frac{\partial h_R}{\partial\tau} + 2(1-\tau)\frac{h_0}{(h_0+h_R)^3}\Big(\frac{\partial h_R}{\partial\tau}\Big)^2 \\ &\qquad - (1-\tau)\frac{\mu h_0}{(h_0+h_R)^2}\frac{\partial^2 h_R}{\partial\tau^2}\Big]\end{aligned}$$

Equation (18) writes as

$$\begin{aligned}
&\frac{\partial^2 W}{\partial M \partial \tau} = \\
&- p(1-\tau)\frac{\mu h_0}{(h_0+h_R)^2}\frac{\partial h_R}{\partial M}\Big(- \frac{\mu h_R}{h_0+h_R} + (1-\tau)\frac{\mu h_0}{(h_0+h_R)^2}\frac{\partial h_R}{\partial \tau}\Big) \\
&- p\Big(\mu\eta - (1-\tau)\frac{\mu h_R}{h_0+h_R}\Big)\Big[\frac{\mu h_0}{(h_0+h_R)^2}\frac{\partial h_R}{\partial M} + 2(1-\tau)\frac{\mu h_0}{(h_0+h_R)^3}\frac{\partial h_R}{\partial M}\frac{\partial h_R}{\partial \tau} \\
&\qquad\qquad - (1-\tau)\frac{\mu h_0}{(h_0+h_R)^2}\frac{\partial^2 h_R}{\partial M \partial \tau}\Big]
\end{aligned}$$

Let us note that rearranging (19) yields that

$$\begin{aligned}
&\Big[\Big(\mu\eta - \frac{\mu h_R}{h_0+h_R}\Big) - \frac{M}{1-\tau}\Big]\Big[\frac{\mu h_0}{(h_0+h_R)^2}\frac{\partial h_R}{\partial M} + 2(1-\tau)\frac{\mu h_0}{(h_0+h_R)^3}\frac{\partial h_R}{\partial M}\frac{\partial h_R}{\partial \tau} \\
&\qquad\qquad - (1-\tau)\frac{\mu h_0}{(h_0+h_R)^2}\frac{\partial^2 h_R}{\partial M \partial \tau}\Big] \\
&= -\frac{\mu h_0}{(h_0+h_R)^2}\frac{\partial h_R}{\partial \tau} - \frac{\mu h_0}{(h_0+h_R)^2}\Big[(1-\tau)\frac{\mu h_0}{(h_0+h_R)^2}\frac{\partial h_R}{\partial \tau} + \frac{M}{1-\tau}\Big]\frac{\partial h_R}{\partial M}.
\end{aligned}$$

Using previous computations, both terms on the RHS are strictly negative. Therefore,

$$\frac{\mu h_0}{(h_0+h_R)^2}\frac{\partial h_R}{\partial M} + 2(1-\tau)\frac{\mu h_0}{(h_0+h_R)^3}\frac{\partial h_R}{\partial M}\frac{\partial h_R}{\partial \tau} - (1-\tau)\frac{\mu h_0}{(h_0+h_R)^2}\frac{\partial^2 h_R}{\partial M \partial \tau} > 0,$$

and thus, using again previous computations,[70]

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

Displacement. Suppose $\eta > \chi$. By Lemma 3, $h_R/(h_0+h_R) = \eta$ and

$$(1-\tau)\Big(\mu\eta - \frac{1}{2}\frac{\mu h_P}{h_0+h_P}\Big)\frac{\mu h_P}{h_0+h_P} = \frac{1}{2}(1-\tau)(\mu\eta)^2 - \mu\Big(\frac{h_P}{h_0+h_P} - \chi\Big)M.$$

Therefore,

$$\frac{\partial h_R}{\partial \tau} = \frac{\partial h_R}{\partial M} = 0,$$

[70] In particular, from (20),

$$-\frac{\mu h_R}{h_0+h_R} + (1-\tau)\frac{\mu h_0}{(h_0+h_R)^2}\frac{\partial h_R}{\partial \tau} < 0.$$

while

$$
\begin{aligned}
&\left[(1-\tau)\Big(\mu\eta - \frac{\mu h_P}{h_0+h_P}\Big)\frac{\mu h_0}{(h_0+h_P)^2} + \frac{\mu h_0}{(h_0+h_P)^2}M\right]\frac{\partial h_P}{\partial \tau} \\
&\qquad = \Big(\mu\eta - \frac{1}{2}\frac{\mu h_P}{h_0+h_P}\Big)\frac{\mu h_P}{h_0+h_P} - \frac{1}{2}(\mu\eta)^2 = -\frac{1}{2}\Big(\frac{\mu h_P}{h_0+h_P} - \mu\eta\Big)^2
\end{aligned}
$$

and

$$
\begin{aligned}
&\left[(1-\tau)\Big(\mu\eta - \frac{\mu h_P}{h_0+h_P}\Big)\frac{\mu h_0}{(h_0+h_P)^2} + \frac{\mu h_0}{(h_0+h_P)^2}M\right]\frac{\partial h_P}{\partial M} \\
&\qquad = -\mu\Big(\frac{h_P}{h_0+h_P} - \chi\Big) = -\frac{1-\tau}{2M}\Big(\frac{\mu h_P}{h_0+h_P} - \mu\eta\Big)^2,
\end{aligned}
$$

where the inequality follows from Lemma 3. Hence,

$$
\frac{\partial h_P}{\partial M} = \frac{1-\tau}{M}\frac{\partial h_P}{\partial \tau} \tag{21}
$$

Moreover,

$$
\begin{aligned}
&\left[-(1-\tau)\Big(\frac{\mu h_0}{(h_0+h_P)^2}\Big)^2 - 2(1-\tau)\Big(\mu\eta - \frac{\mu h_P}{h_0+h_P}\Big)\frac{\mu h_0}{(h_0+h_P)^3} - \frac{2\mu h_0}{(h_0+h_P)^3}M\right]\frac{\partial h_P}{\partial M}\frac{\partial h_P}{\partial \tau} \\
&+ \left[(1-\tau)\Big(\mu\eta - \frac{\mu h_P}{h_0+h_P}\Big)\frac{\mu h_0}{(h_0+h_P)^2} + \frac{\mu h_0}{(h_0+h_P)^2}M\right]\frac{\partial^2 h_P}{\partial M \partial \tau} \\
&\qquad = -\frac{\mu h_0}{(h_0+h_P)^2}\frac{\partial h_P}{\partial \tau} + \Big(\mu\eta - \frac{\mu h_P}{h_0+h_P}\Big)\frac{\mu h_0}{(h_0+h_P)^2}\frac{\partial h_P}{\partial M}
\end{aligned} \tag{22}
$$

Hence, in particular, (16) writes as

$$
\begin{aligned}
\frac{\partial W}{\partial \tau} &= -p\tau(\mu\eta)^2 + (1-p)\Big(\mu\eta - (1-\tau)\frac{\mu h_P}{h_0+h_P}\Big)\Big(-\frac{\mu h_P}{h_0+h_P} + (1-\tau)\frac{\mu h_0}{(h_0+h_P)^2}\frac{\partial h_P}{\partial \tau}\Big) \\
&= -p\tau(\mu\eta)^2 + (1-p)\Big(\mu\eta - (1-\tau)\frac{\mu h_P}{h_0+h_P}\Big)\frac{(1-\tau)\Big(\Big(\frac{\mu h_P}{h_0+h_P}\Big)^2 - (\mu\eta)^2\Big) - \frac{\mu h_P}{h_0+h_P}M}{(1-\tau)\Big(\mu\eta - \frac{\mu h_P}{h_0+h_P}\Big) + M}
\end{aligned} \tag{23}
$$

By Lemma 3, $\mu\eta > \mu h_P/(h_0+h_P)$. Therefore, $\partial h_P/\partial\tau < 0$, and from (23), any solution τ^* to $\partial W(\tau)/\partial\tau = 0$ is necessarily strictly negative: $\tau^* < 0$, and such that $\mu\eta > (1-$

$\tau^*)\mu h_P/(h_0 + h_P)$.[71]

Equation (18) writes as

$$\begin{aligned}
&\frac{\partial^2 W}{\partial M \partial \tau} = \\
&- (1-p)(1-\tau)\frac{\mu h_0}{(h_0+h_P)^2}\frac{\partial h_P}{\partial M}\bigg(-\frac{\mu h_P}{h_0+h_P} + (1-\tau)\frac{\mu h_0}{(h_0+h_P)^2}\frac{\partial h_P}{\partial \tau}\bigg) \\
&- (1-p)\bigg(\mu\eta - (1-\tau)\frac{\mu h_P}{h_0+h_P}\bigg)\bigg[\frac{\mu h_0}{(h_0+h_P)^2}\frac{\partial h_P}{\partial M} + 2(1-\tau)\frac{\mu h_0}{(h_0+h_P)^3}\frac{\partial h_P}{\partial M}\frac{\partial h_P}{\partial \tau} \\
&\qquad - (1-\tau)\frac{\mu h_0}{(h_0+h_P)^2}\frac{\partial^2 h_P}{\partial M \partial \tau}\bigg]
\end{aligned}$$

Let us note that rearranging (22) yields that

$$\begin{aligned}
&\bigg[\bigg(\mu\eta - \frac{\mu h_P}{h_0+h_P}\bigg) + \frac{M}{1-\tau}\bigg]\bigg[\frac{\mu h_0}{(h_0+h_P)^2}\frac{\partial h_P}{\partial M} + 2(1-\tau)\frac{\mu h_0}{(h_0+h_P)^3}\frac{\partial h_P}{\partial M}\frac{\partial h_P}{\partial \tau} \\
&\qquad - (1-\tau)\frac{\mu h_0}{(h_0+h_P)^2}\frac{\partial^2 h_P}{\partial M \partial \tau}\bigg] \\
&= \frac{\mu h_0}{(h_0+h_P)^2}\frac{\partial h_P}{\partial \tau} + \frac{\mu h_0}{(h_0+h_P)^2}\bigg[-(1-\tau)\frac{\mu h_0}{(h_0+h_P)^2}\frac{\partial h_P}{\partial \tau} + \frac{M}{1-\tau}\bigg]\frac{\partial h_P}{\partial M} \\
&= \frac{\mu h_0}{(h_0+h_P)^2}\frac{M}{1-\tau}\frac{\partial h_P}{\partial M} + \frac{\mu h_0}{(h_0+h_P)^2}\bigg[-(1-\tau)\frac{\mu h_0}{(h_0+h_P)^2}\frac{\partial h_P}{\partial \tau} + \frac{M}{1-\tau}\bigg]\frac{\partial h_P}{\partial M}
\end{aligned}$$

---

[71] Equation (17) writes as

$$\begin{aligned}
&\frac{\partial^2 W}{\partial \tau^2} \\
&= -p(\mu\eta)^2 - (1-p)\bigg(\frac{\mu h_P}{h_0+h_P} - (1-\tau)\frac{\mu h_0}{(h_0+h_P)^2}\frac{\partial h_P}{\partial \tau}\bigg)^2 \\
&\quad - (1-p)\bigg(\mu\eta - (1-\tau)\frac{\mu h_P}{h_0+h_P}\bigg)\bigg[\frac{2\mu h_0}{(h_0+h_P)^2}\frac{\partial h_P}{\partial \tau} + 2(1-\tau)\frac{h_0}{(h_0+h_P)^3}\bigg(\frac{\partial h_P}{\partial \tau}\bigg)^2 \\
&\qquad - (1-\tau)\frac{\mu h_0}{(h_0+h_P)^2}\frac{\partial^2 h_P}{\partial \tau^2}\bigg]
\end{aligned}$$

and thus in particular, for τ^* such that $W(\tau^*) = 0$,

$$\begin{aligned}
&\frac{\partial^2 W}{\partial \tau^2} \\
&= -p(\mu\eta)^2 - \frac{p^2}{1-p}\left(\frac{\tau(\mu\eta)^2}{\mu\eta - (1-\tau)\frac{\mu h_P}{h_0+h_P}}\right)^2 \\
&\quad - p\bigg(\mu\eta - (1-\tau)\frac{\mu h_P}{h_0+h_P}\bigg)\bigg[\frac{2\mu h_0}{(h_0+h_P)^2}\frac{\partial h_P}{\partial \tau} + 2(1-\tau)\frac{h_0}{(h_0+h_P)^3}\bigg(\frac{\partial h_P}{\partial \tau}\bigg)^2 \\
&\qquad - (1-\tau)\frac{\mu h_0}{(h_0+h_P)^2}\frac{\partial^2 h_P}{\partial \tau^2}\bigg]
\end{aligned}$$

where the last equality follows from (21). Consequently, using again previous computations,[72]

$$
\begin{aligned}
&\frac{\left(\mu\eta - \frac{\mu h_P}{h_0+h_P}\right) + \frac{M}{1-\tau}}{(1-p)\frac{\mu h_0}{(h_0+h_P)^2}\frac{\partial h_P}{\partial M}}\frac{\partial^2 W}{\partial M \partial \tau} \\
&= -\left[(1-\tau)\left(\left(\frac{\mu h_P}{h_0+h_P}\right)^2 - (\mu\eta)^2\right) - \frac{\mu h_P}{h_0+h_P}M\right] \\
&\quad - \left(\mu\eta - (1-\tau)\frac{\mu h_P}{h_0+h_P}\right)\left[\frac{2M}{1-\tau} - (1-\tau)\frac{\mu h_0}{(h_0+h_P)^2}\frac{\partial h_P}{\partial \tau}\right] \\
&= \left(\frac{3\mu h_P}{h_0+h_P} - \frac{2\mu\eta}{1-\tau}\right)M - (1-\tau)\left(\left(\frac{\mu h_P}{h_0+h_P}\right)^2 - (\mu\eta)^2\right) \\
&\quad - (1-\tau)\left(\mu\eta - (1-\tau)\frac{\mu h_P}{h_0+h_P}\right)\frac{\frac{1}{2}\left(\frac{\mu h_P}{h_0+h_P} - \mu\eta\right)^2}{(1-\tau)\left(\mu\eta - \frac{\mu h_P}{h_0+h_P}\right) + M} \\
&= \left(\frac{3\mu h_P}{h_0+h_P} - \frac{2\mu\eta}{1-\tau}\right)M \\
&\quad + (1-\tau)\left(\mu\eta - \frac{\mu h_P}{h_0+h_P}\right)\left[\mu\eta + \frac{\mu h_P}{h_0+h_P} - \frac{1}{2}\left(\mu\eta - (1-\tau)\frac{\mu h_P}{h_0+h_P}\right)\frac{\mu\eta - \frac{\mu h_P}{h_0+h_P}}{(1-\tau)\left(\mu\eta - \frac{\mu h_P}{h_0+h_P}\right) + M}\right]
\end{aligned}
$$

The second term on the RHS is strictly positive for any $\tau < 0$ (as $\eta > h_P/(h_0+h_P)$).[73]

Fixing $\tau \leq 0$, h_P strictly decreases with M, and is such that $h_P/(h_0+h_P) = \eta$ for $M = 0$. Hence, for any $\tau \leq 0$, there exists $\underline{M}(\tau) > 0$, with $\underline{M}$ strictly increasing with τ, such that for any $M < \underline{M}(\tau)$ and any $\tau' \leq \tau$, the first term on the RHS is strictly positive for any $\tau' \leq \tau$. In particular, for any $M < \underline{M}(0)$, the first term on the RHS is strictly positive for any $\tau \leq 0$.

[72] In particular, as in (23),

$$
-\frac{\mu h_P}{h_0+h_P} + (1-\tau)\frac{\mu h_0}{(h_0+h_P)^2}\frac{\partial h_P}{\partial \tau} = \frac{(1-\tau)\left[\left(\frac{\mu h_P}{h_0+h_P}\right)^2 - (\mu\eta)^2\right] - \frac{\mu h_P}{h_0+h_P}M}{(1-\tau)\left(\mu\eta - \frac{\mu h_P}{h_0+h_P}\right) + M} < 0.
$$

[73] Indeed, for any $\tau < 0$,

$$
\begin{aligned}
&\mu\eta + \frac{\mu h_P}{h_0+h_P} - \frac{1}{2}\left(\mu\eta - (1-\tau)\frac{\mu h_P}{h_0+h_P}\right)\frac{\mu\eta - \frac{\mu h_P}{h_0+h_P}}{(1-\tau)\left(\mu\eta - \frac{\mu h_P}{h_0+h_P}\right) + M} \\
&> \mu\eta + \frac{\mu h_P}{h_0+h_P} - \frac{1}{2}\left(\mu\eta - \frac{\mu h_P}{h_0+h_P}\right) = \frac{1}{2}\left(\mu\eta + \frac{3\mu h_P}{h_0+h_P}\right).
\end{aligned}
$$

Therefore, for any $M < \underline{M}(0)$,

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

H Proof of Proposition 6

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 headstart w conditional on observables I.

Let us start with a general observation. In a separating equilibrium, for any choice of period-1 activity h_1 and effort e_1, any choice of period-2 activity h_2, any realization of period-1 performance y_1, and any audience's beliefs $\hat{e}_1, \hat{e}_2$, an agent's optimal effort in period-2, $e_2^*(h_1, e_1, y_1, h_2)$ solves

$$\max_e \frac{\mu h_2}{h_0 + h_1 + h_2} e - g(e),$$

and thus e_2^* is (uniquely) given by

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

Let us now look for an equilibrium in pure strategies. Suppose first that there is no headstart 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

$$\max_{h_2} \Bigg(\frac{\mu h_1}{h_0 + h_1 + h_2} [y_1 - \mathbb{E}[\hat{e}_1(h_1)]] + \mu\eta\hat{e}_1(h_1) \\ + \frac{\mu h_2}{h_0 + h_1 + h_2} \frac{h_1}{h_0 + h_1} [y_1 - e_1] + \mu\eta e_2^*(h_1, h_2) - g(e_2^*(h_1, h_2)) \Bigg)$$

i.e. to

$$\max_{h_2}\left(\frac{\mu h_1}{h_0+h_1+h_2}[e_1-\hat{e}_1(h_1)]+\mu\eta e_2^*(h_1,h_2)-g(e_2^*(h_1,h_2))\right) \tag{24}$$

A solution $h_2^\dagger$ exists if and only if $\mu\eta+[h_1/(h_0+h_1)][\hat{e}_1(h_1)-e_1]<1$. Then, the objective being continuously differentiable, first strictly increasing then strictly decreasing 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

$$\begin{aligned}\max_{e_1}\mathbb{E}_{\theta+\varepsilon_1}\Bigg[&\frac{\mu h_1}{h_0+h_1+h_2^\dagger}[\theta+\varepsilon_1+e_1-\hat{e}_1(h_1)]+\mu\eta\hat{e}_1(h_1)-\delta^{-1}g(e_1)\\&+\frac{\mu h_2^\dagger}{h_0+h_1+h_2^\dagger}\frac{h_1}{h_0+h_1}[\theta+\varepsilon_1]+\mu\eta e_2^*(h_1,h_2^\dagger)-g(e_2^*(h_1,h_2^\dagger))\Bigg]\end{aligned}$$

i.e. to

$$=\max_{e_1}\left(\frac{\mu h_1}{h_0+h_1+h_2^\dagger}[e_1-\hat{e}_1(h_1)]-\delta^{-1}g(e_1)+\mu\eta e_2^*(h_1,h_2^\dagger)-g(e_2^*(h_1,h_2^\dagger))\right) \tag{25}$$

as neither $\hat{e}_1(h_1)$ nor $h_2^\dagger$ depend on the realization of $\theta+\varepsilon_1$.[74] As $h_2^\dagger$ is a solution to (24), the agent's period-1 effort $e_1^\dagger$ is thus (uniquely) given by

$$e_1^\dagger(h_1,\hat{e}(h_1))=\frac{\delta\mu h_1}{h_0+h_1+h_2^\dagger(h_1,e_1,\hat{e}(h_1))}$$

[74]For $e_1>0$, the first derivative of the objective with respect to e_1 is equal to

$$\begin{aligned}&\frac{\mu h_1}{h_0+h_1+h_2^\dagger}-\delta^{-1}g'(e_1)+\left(\frac{\mu h_1}{(h_0+h_1+h_2^\dagger)^2}\left[\hat{e}_1(h_1)-e_1\right]+\left[\mu\eta-g'(e_2^*(h_1,h_2^\dagger))\right]\frac{\partial e_2^*}{\partial h_2}\right)\frac{\partial h_2^\dagger}{\partial e_1}\\=&\frac{\mu h_1}{h_0+h_1+h_2^\dagger}-\delta^{-1}g'(e_1),\end{aligned}$$

as $h_2^\dagger$ is a solution to (24), while the second derivative of the objective with respect to e_1 is thus equal to

$$-\frac{\mu h_1}{(h_0+h_1+h_2^\dagger)^2}\frac{\partial h_2^\dagger}{\partial e_1}-\delta^{-1}g''(e_1)=\left(\frac{h_1}{h_0+h_1}\right)^2-\delta^{-1}<0.$$

A necessary condition for equilibrium 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 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 solution to (24) 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:

$$h_2^\dagger\left(\mu(1-\eta) - \frac{h_1}{h_0 + h_1}[\hat{e}_1(h_1) - e_1]\right) = (h_0 + h_1)\left(\mu\eta + \frac{h_1}{h_0 + h_1}[\hat{e}_1(h_1) - e_1]\right)$$

And thus

$$\begin{aligned}
e_1^\dagger &= \frac{\delta\mu h_1}{h_0+h_1+h_2^\dagger(h_1,e_1^\dagger,\hat{e}(h_1))} \\
&= \frac{\delta\mu h_1}{h_0+h_1}\frac{\mu(1-\eta)-\frac{h_1}{h_0+h_1}[\hat{e}_1(h_1)-e_1]}{\mu(1-\eta)-\frac{h_1}{h_0+h_1}[\hat{e}_1(h_1)-e_1]+\mu\eta+\frac{h_1}{h_0+h_1}[\hat{e}_1(h_1)-e_1]} \\
&= \frac{\delta\mu h_1}{h_0+h_1}\frac{\mu(1-\eta)-\frac{h_1}{h_0+h_1}[\hat{e}_1(h_1)-e_1^\dagger]}{\mu}
\end{aligned}$$

i.e. rearranging and replacing $\hat{e}_1(h_1)$ by its explicit expression,

$$\left[1-\delta\Big(\frac{h_1}{h_0+h_1}\Big)^2\right]e_1^\dagger = \delta\mu(1-\eta)\frac{\delta\mu h_1}{h_0+h_1}\left[1-\delta\Big(\frac{h_1}{h_0+h_1}\Big)^2\right],$$

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

$$\mu\eta+\frac{h_1}{h_0+h_1}[\hat{e}_1(h_1)-e_1^\dagger]=\mu\eta\Big(1-\frac{\delta h_1^2}{(h_0+h_1)^2}\Big)>0,$$

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 her 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^*.

Let us now introduce headstart inequality. We look for a separating equilibrium in pure strategies, i.e. we look for equilibria in which the agents' headstart is perfectly revealed. As a consequence, in any such equilibrium, the audience's expectation of the agent's headstart does not depend on realized performances y_1, y_2. and thus $\hat{w}$ does not depend on performances y_1, y_2. We denote by $\hat{w}(h_1, h_2)$ the audience's expectation of the agent's headstart after observing h_1, h_2.

In a separating equilibrium, the agent's period-2 activity choice is a solution, if any, to

$$\begin{aligned}
&\max_{h_2} \Bigg(\frac{\mu h_1}{h_0 + h_1 + h_2}[y_1 - \hat{e}_1(h_1) - \hat{w}(h_1, h_2)] + \mu\eta\hat{e}_1(h_1) + \mu\chi\hat{w}(h_1, h_2) \\
&\qquad + \frac{\mu h_2}{h_0 + h_1 + h_2}\frac{h_1}{h_0 + h_1}[y_1 - e_1 - \hat{w}(h_1, h_2)] + \mu\eta e_2^*(h_1, h_2) - g(e_2^*(h_1, h_2)) \Bigg) \\
= &\max_{h_2} \Bigg(- \frac{\mu h_1}{h_0 + h_1 + h_2}\hat{e}_1(h_1) - \frac{\mu h_2}{h_0 + h_1 + h_2}\frac{h_1}{h_0 + h_1} e_1 \\
&\qquad - \mu\left(\frac{h_1}{h_0 + h_1} - \chi\right)\hat{w}(h_1, h_2) + \mu\eta e_2^*(h_1, h_2) - g(e_2^*(h_1, h_2)) \Bigg)
\end{aligned}$$

Hence, the agent's objective does not depend on her actual headstart w. As mentioned in the text, we then assume that in such a case, the audience's equilibrium belief about the agent's headstart does not depend on her choice of period-2 activity h_2. As a consequence, the agent chooses h_2 by solving the same maximization program (24) as without headstart inequality, and in equilibrium, separation must occur in the first period.

Hence, the agent's period-1 effort is a solution, if any, to

$$\begin{aligned}
&\max_{e_1} \mathbb{E}_{\theta+\varepsilon_1}\Big[\frac{\mu h_1}{h_0+h_1+h_2^\dagger}[\theta+\varepsilon_1+e_1+w-\hat{e}_1(h_1)-\hat{w}(h_1)] \\
&\qquad + \frac{\mu h_2^\dagger}{h_0+h_1+h_2^\dagger}\frac{h_1}{h_0+h_1}[\theta+\varepsilon_1+w-\hat{w}(h_1)] \\
&\qquad + \mu\eta\hat{e}_1(h_1)-\delta^{-1}g(e_1)+\mu\chi\hat{w}(h_1)+\mu\eta e_2^*(h_1,h_2^\dagger)-g(e_2^*(h_1,h_2^\dagger))\Big] \\
&= \max_{e_1} \mathbb{E}_{\theta+\varepsilon_1}\Big[\frac{\mu h_1}{h_0+h_1+h_2^\dagger}[e_1-\hat{e}_1(h_1)]-\mu\Big(\frac{h_1}{h_0+h_1}-\chi\Big)\hat{w}(h_1) \\
&\qquad + \mu\eta\hat{e}_1(h_1)-\delta^{-1}g(e_1)+\mu\eta e_2^*(h_1,h_2^\dagger)-g(e_2^*(h_1,h_2^\dagger))\Big] \\
&= \max_{e_1}\Big(\frac{\mu h_1}{h_0+h_1+h_2^\dagger}[e_1-\hat{e}_1(h_1)]-\delta^{-1}g(e_1)+\mu\eta e_2^*(h_1,h_2^\dagger)-g(e_2^*(h_1,h_2^\dagger))\Big)
\end{aligned}$$

which is the same maximization program (25) as without headstart inequality.

The same arguments as in the case of no headstart inequality then imply that the strategies $e_1^*(h_1)$ and $h_2^*(h_1)$ and (degenerate) beliefs $\hat{e}_1(h_1) = e_1^*(h_1)$ defined above, form the unique equilibrium in pure strategies of the continuation game starting after period-1 activity choice: as in the absence of headstart inequality, 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.

Lastly, the agent's choice of h_1 depends on her actual headstart: it is given by the solution, if any, to

$$\begin{aligned}
\max_{h_1} \Big[&\mu\eta e_1^*(h_1)-\delta^{-1}g(e_1^*(h_1))+\mu\eta e_2^*(h_1,h_2^*(h_1))-g(e_2^*(h_1,h_2^*(h_1))) \\
&+\frac{\mu(h_1+h_2^*(h_1))}{h_0+h_1+h_2^*(h_1)}w-\mu\Big(\frac{(h_1+h_2^*(h_1)}{h_0+h_1+h_2^*(h_1)}-\chi\Big)\hat{w}(h_1)\Big],
\end{aligned}$$

i.e. 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 the solution, if any, to

$$\max_{h_1}\Big[\mu\eta e_1^*(h_1)-\delta^{-1}g(e_1^*(h_1))+\frac{\mu(h_1+h_2^*(h_1))}{h_0+h_1+h_2^*(h_1)}w-\mu\Big(\frac{h_1+h_2^*(h_1)}{h_0+h_1+h_2^*(h_1)}-\chi\Big)\hat{w}(h_1)\Big].$$

The same arguments as in the proof of Proposition 1 (see Appendix C) then yield the result.

I Proofs of Proposition 7 and Corollary 3

We look for separating equilibria. 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

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

as the weights on within- and across-activity images sum to 1.

For each $k \in \{1, ..., N\}$, let $U_k(\zeta) \equiv \beta_k + \zeta\mu\eta e^*(h_k) - g(e^*(h_k))$.[75] With relative image concerns, in a separating equilibrium, each agent with headstart w chooses her activity by solving:

$$\max_{k \in \{1,...,N\}} \left(U_k(\zeta) + \frac{\mu h_k}{h_0 + h_k} w - \frac{\mu h_k}{h_0 + h_k} \mathbb{E}[w|k] + \zeta\mu\chi\mathbb{E}[w|k] \right)$$

i.e. by solving (P), only replacing η by $\zeta\eta$ and χ by $\zeta\chi$. Proposition 7 and Corollary 3 then follows from the proofs of Proposition 1 (see Appendix C) and comparative statics with respect to ζ.

[75] Similarly, in the case of a continuum of activities, with Assumption 1, for each $h \in \mathbb{R}_+$, let $U(H, \zeta) \equiv \beta(e^*(h)) + \zeta\mu\eta e^*(h) - g(e^*(h))$. Each agent with headtart w then chooses h to maximize

$$U(h, \zeta) + \frac{\mu h}{h_0 + h} w - \frac{\mu h}{h_0 + h} \mathbb{E}[w|h] + \zeta\mu\chi\mathbb{E}[w|h]$$