Online Appendix for

"Cooptation: Meritocracy vs. Homophily in Organizations"

Paul-Henri Moisson Jean Tirole

May 15, 2020

A Value functions for majority and minority members

Value function for a majority member. Under restrictions (i)-(ii) on the majority's strategy, let V_i^- denote the expected value function conditional on the minority candidate being more talented, and V_i^+ denote the expected value function conditional on the complementary event. The value function for a majority member (see below for that of a minority member) writes for any $k \leq M \leq N-1$[1],

$$V_M = xV_M^- + (1-x)V_M^+ \tag{2}$$

$$\text{where} \quad \begin{cases} V_M^- = \max\Big\{b + \delta\Big(\dfrac{M}{N-1}V_M + (1-\dfrac{M}{N-1})V_{M+1}\Big), \\ \qquad\qquad s + \delta\Big(\dfrac{M-1}{N-1}V_{M-1} + (1-\dfrac{M-1}{N-1})V_M\Big)\Big\} \\ V_M^+ = b + \dfrac{\overline{x}}{1-x}s + \delta\Big(\dfrac{M}{N-1}V_M + \Big(1-\dfrac{M}{N-1}\Big)V_{M+1}\Big) \end{cases}$$

With probability x, the majority faces a trade-off between selecting a talented minority member (yielding payoff s) and picking the less talented majority candidate (yielding payoff b). With probability $1-x$, the choice is a no-brainer and the majority candidate brings average payoff $b + \overline{x}s/(1-x)$ where $\overline{x}/(1-x)$ is the conditional probability of that candidate's being talented. Furthermore, for a majority member, recruiting a majority candidate when the majority has size M in period t yields an *ex-ante* (i.e. before departure) majority size of $M+1$. Three events might then happen at the beginning of period $t+1$ before the vote takes place: *(i)* with probability $1/N$ (which is already embodied in the discount factor δ),

[1] Equation (2) applies even when $M = N-1$ as the majority size $M+1$ becomes irrelevant (its probability being nil).

the majority member quits the organization, which gives him zero payoff; *(ii)* with probability M/N, *another* majority member quits, and thus the majority size decreases to M; *(iii)* with probability $(N-M-1)/N$, a minority member quits, and thus the majority size remains equal to $M+1$.

Value function for a minority member. If the majority recruits the majority candidate in period t, then at the beginning of period $t+1$: *(i)* with probability $1/N$, the minority member quits the organization, which gives her zero payoff; *(ii)* with probability $(M+1)/N$, a majority member quits, and thus the majority size decreases to M; *(iii)* with probability $(N-M-2)/N$, *another* minority member quits, and thus the majority size remains equal to $M+1$.

Let $\sigma(M) \in \{0,1\}$ be defined for any $k \leq M \leq N-1$ as

$$\sigma(M) = \begin{cases} 1 & \text{if } b + \delta\Big(\dfrac{M}{N-1}V_M + (1-\dfrac{M}{N-1})V_{M+1}\Big) > s + \delta\Big(\dfrac{M-1}{N-1}V_{M-1} + (1-\dfrac{M-1}{N-1})V_M\Big) \\ 0 & \text{otherwise.} \end{cases}$$

The value function for a minority member writes for any $k \leq M \leq N-2$:

$$V_{N-M-1} = xV^-_{N-M-1} + (1-x)V^+_{N-M-1} \tag{3}$$

$$\text{where} \quad \begin{cases} V^-_{N-M-1} = \sigma(M)\delta\Big(\dfrac{M+1}{N-1}V_{N-M-1} + (1-\dfrac{M+1}{N-1})V_{N-M-2}\Big) \\ \qquad\qquad +(1-\sigma(M))\Big[s + b + \delta\Big(\dfrac{M}{N-1}V_{N-M} + (1-\dfrac{M}{N-1})V_{N-M-1}\Big)\Big] \\ V^+_{N-M-1} = \dfrac{\overline{x}}{1-x}s + \delta\Big(\dfrac{M+1}{N-1}V_{N-M-1} + \Big(1-\dfrac{M+1}{N-1}\Big)V_{N-M-2}\Big) \end{cases}$$

B Proof of Lemma 1

The result for $N = 4$ derives from straightforward computations[2]. We assume in the following $N \geq 6$. Assume $s > b > 0$ and $x \in (0, 1/2)$.

[2] Using (2) and (3), for the entrenched equilibrium, one has

$$\Big[1 - \frac{2\delta}{3}(1-x)\Big](V_3^{\mathrm{e}} - V_2^{\mathrm{e}}) = x(s-b)$$

and thus $V_1^{\mathrm{e}} = \overline{x}s/(1-\delta) < (b+\overline{x}s)/(1-\delta) < V_2^{\mathrm{e}} < V_3^{\mathrm{e}}$. Similarly for the meritocratic equilibrium:

$$\begin{cases} \Big[1 - \dfrac{x\delta}{3} - \dfrac{2\delta}{3}(1-x)\Big](V_3^{\mathrm{m}} - V_2^{\mathrm{m}}) = \dfrac{x\delta}{3}(V_2^{\mathrm{m}} - V_1^{\mathrm{m}}) \\ \Big[1 - \delta(1-x)\Big](V_2^{\mathrm{m}} - V_1^{\mathrm{m}}) = (1-2x)b + \delta\dfrac{(1-x)}{3}(V_3^{\mathrm{m}} - V_2^{\mathrm{m}}) \end{cases}$$

and thus $V_1^{\mathrm{m}} < V_2^{\mathrm{m}} < V_3^{\mathrm{m}}$, and $V_2^{\mathrm{m}} - V_1^{\mathrm{m}} > V_3^{\mathrm{m}} - V_1^{\mathrm{m}}$.

Proof of (i). Consider, first, the entrenched equilibrium. We omit the e superscript in order to alleviate the notation. For any $M \in \{k-1, ..., N-2\}$, let $u_M \equiv V_{M+1} - V_M$. By writing the expression of the value function from (2) in $M \in \{k+1, ..., N-1\}$ (thus writing V_M as a function of V_{M-1}, V_M and V_{M+1}), and then substracting the expression in M from the expression in $M+1$ yields for any $M \in \{k+1, ..., N-2\}$[3], and rearranging, we get that

$$\left[1 - \delta x\left(1 - \frac{M}{N-1}\right) - \delta(1-x)\frac{M}{N-1}\right]u_M = \delta x\frac{M-1}{N-1}u_{M-1} + \delta(1-x)\left(1 - \frac{M+1}{N-1}\right)u_{M+1} \tag{4}$$

We show the result by contradiction and by induction. Suppose $u_{N-2} \leq 0$. Then Equation (4) for $M = N-2$ implies

$$\left[1 - \delta\frac{x}{N-1} - \delta(1-x)\frac{N-2}{N-1}\right]u_{N-2} = \delta x\frac{N-3}{N-1}u_{N-3}$$

Therefore, $u_{N-3} \leq 0$ and $u_{N-3} \leq u_{N-2}$. We then proceed by induction to show that for any $M \in \{k+1, ..., N-2\}$, $u_{M-1} \leq u_M \leq 0$. Assume the result holds for all indices in $\{M+1, ..., N-2\}$. Then Equation (4) implies

$$\left[1 - \delta x\left(1 - \frac{M}{N-1}\right) - \delta(1-x)\frac{M}{N-1}\right]u_M \geq \delta x\frac{M-1}{N-1}u_{M-1} + \delta(1-x)\left(1 - \frac{M+1}{N-1}\right)u_M$$

$$\text{i.e.} \quad \left[1 - \delta x\left(1 - \frac{M}{N-1}\right) - \delta(1-x)\frac{N-2}{N-1}\right]u_M \geq \delta x\frac{M-1}{N-1}u_{M-1}$$

Consequently, $u_{M-1} \leq u_M \leq 0$. Hence the result by induction. In particular, one has $u_k \leq 0$, i.e. $V_{k+1} - V_k \leq 0$. However, writing Equation (2) in $k+1$ and k and taking the difference yields

$$u_k = x(s-b) + \delta(1-x)\left[\left(1 - \frac{k+1}{N-1}\right)u_{k+1} + \frac{k}{N-1}u_k\right] \tag{5}$$

and thus

$$0 \geq \left[1 - \delta(1-x)\frac{N-2}{N-1}\right]u_k \geq x(s-b) > 0,$$

which is a contradiction. Therefore $V_{N-1} - V_{N-2} = u_{N-2} > 0$. It is then easy to see that the same induction argument used above, using repeatedly Equation (4), shows that $u_{M-1} > u_M > 0$ for any $M \in \{k+1, ..., N-2\}$. Hence the result for the entrenched equilibrium.

Consider now the meritocratic equilibrium. We again omit the superscript on the value

[3] With the abuse of notation $u_{N-1} = 0$, which is irrelevant since the coefficient of u_{N-1} is nil.

function in order to alleviate the notation. Let $u_i \equiv V_{i+1} - V_i$ for any $i \in \{1, ..., N-2\}$. Note that Equation (4) holds for any $M \in \{k, ..., N-2\}$. The argument is similar to the one used in the entrenched equilibrium: the idea is again to suppose that the majority prefers being in $N-2$ over $N-1$ and reach a contradiction. Note that there are two differences with respect to the entrenchment setup, as *(a)* the contradiction stems from the loss of homophily payoff (when candidates have same talent) associated with losing the majority, and *(b)* since there is no entrenchment and majorities switch, in order to reach the contradiction, the induction needs to go down until a group size of 1.

Assume by contradiction that $u_{N-2} \leq 0$. Then, by induction, this implies that for any $M \in \{k, ..., N-2\}$, $u_{M-1} \leq u_M \leq 0$, and thus in particular $u_{k-1} \leq u_k \leq 0$.

Consider now u_1. Writing the expression of the value function from (3) in $M \in \{k+1, ..., N-1\}$ (thus writing V_{N-M-1} as a function of V_{N-M-2}, V_{N-M-1} and V_{N-M}), and then substracting the expression in $N-M-1$ from the expression in $N-M$ (and rearranging) yields for any $M \in \{k+2, ..., N-2\}$:

$$\left[1 - \delta(1-x)\frac{M}{N-1} - \delta x\left(1 - \frac{M}{N-1}\right)\right]u_{N-M-1}$$
$$= \delta(1-x)\left(1 - \frac{M+1}{N-1}\right)u_{N-M-2} + \delta x\frac{M-1}{N-1}u_{N-M} \tag{6}$$

and in particular,

$$\left[1 - \delta\frac{x}{N-1} - \delta(1-x)\frac{N-2}{N-1}\right]u_1 = \delta x\frac{N-3}{N-1}u_2$$

By the usual induction argument using (6), $u_1 > 0$ implies $0 < u_1 < u_2 < ... < u_{k-2} < u_{k-1}$, which contradicts $u_{k-1} \leq 0$. Hence $u_1 \leq 0$ and the same induction argument now implies $0 \geq u_1 \geq u_2 \geq ... \geq u_{k-2} \geq u_{k-1}$.

However, substracting Equation (2) in k and Equation (3) in $k-1$ yields after rearranging:

$$\left[1 - \delta(1-x)\right]u_{k-1} = (1-2x)b + \delta(1-x)\left[\left(1 - \frac{k}{N-1}\right)u_k + \left(1 - \frac{k+1}{N-1}\right)u_{k-2}\right] \tag{7}$$

The contradiction then obtains by summing the above equation together with Equations (4) and (6) over all indexes (and rearranging), which gives:

$$\left(1 - \delta\frac{x}{N-1} - \delta(1-x)\right)(u_1 + u_{N-2}) + (1-\delta)\sum_{i=2}^{N-3} u_i = (1-2x)b > 0$$

This contradicts the fact that $u_i \leq 0$ for all $i \in \{1, ..., N-2\}$. Therefore $u_{N-2} > 0$. Using repeatedly the usual induction argument with Equation (4) yields the result.

The proof of claim *(ii)* relies on the same arguments as the proof of *(i)* and is thus omitted for the sake of brevity.

Claim *(iii)* derives from arguments analogous to the ones used in the proofs of *(i)* and *(ii)*. The result is obvious with *(i)* for the meritocratic equilibrium. The result for the entrenchment equilibrium obtains by considering the sequence $V_i - V_{N-1-i}$ for $i \in \{k, ..., N-2\}$ and using (2)-(3).[4]

Suppose by contradiction that $V_k - V_{k-1} < 0$. This implies that $V_{k+1} - V_{k-2} < V_k - V_{k-1} < 0$, and thus by induction that $V_{N-1} - V_1 < V_{N-2} - V_1 < ... < V_k - V_{k-1} < 0$, which contradicts $V_{N-1} \geq V_{N-2}$ as shown above. (Another contradiction would be reached by summing as above the analogues of (4)-(6) and noting that the RHS is positive whenever $x \leq 1/2$). Hence $V_k - V_{k-1} \geq 0$. If $V_{k+1} - V_{k-2} < 0$, the same contradiction is reached again as then $V_{N-1} - V_1 < V_{N-2} - V_1 < ... < V_{k+1} - V_k < 0$ (Again, one could sum over $i \in \{k+1, ..., N-2\}$ the analogues of (4)-(6) and note that the RHS is positive whenever $x \leq 1/2$). The result obtains by induction: for any $i \in \{k, ..., N-2\}$, $V_i - V_{N-1-i} \geq 0$. Details of the proof show that the inequality is strict if and only if $b > 0$, or $x > 0$ and $s > 0$.

C Proof of Proposition 1

C.1 Proof of Proposition 1-(i)

We show that all symmetric Markov Perfect equilibria in weakly undominated strategies are canonical. To this end, we first define two classes of equilibria depending on $\sigma(k)$, the probability that the organization chooses an inferior in-group candidate when $M = k$:

- k-entrenched equilibria, i.e. equilibria with $\sigma(k) = 1$,
- k-meritocratic equilibria, i.e. equilibria with $\sigma(k) = 0$.

[4] Namely, for $M \in \{k+1, ..., N-3\}$,

$$\left[1 - \delta(1-x)\frac{M}{N-1} - \delta x\left(1 - \frac{M-1}{N-1}\right)\right](V_M - V_{N-M-1}) - (1-2x)b + \frac{\delta}{N-1}\left[(1-x)u_{N-M-2} + xu_{N-M-1}\right]$$
$$= \delta(1-x)\left(1 - \frac{M}{N-1}\right)(V_{M+1} - V_{N-M-2}) + \delta x\frac{M-1}{N-1}(V_{M-1} - V_{N-M})$$

while for $M = k$ and $M = N-2$,

$$\left[1 - \delta\frac{k}{N-1}\right](V_k - V_{k-1}) = b - \frac{\delta}{N-1}u_{k-2} + \delta\left(1 - \frac{k}{N-1}\right)(V_{k+1} - V_{k-2}),$$

$$\left[1 - \delta(1-x)\frac{N-2}{N-1} - \delta x\frac{2}{N-1}\right](V_{N-2} - V_1) = (1-2x)b - \frac{\delta x}{N-1}u_1 + \delta\frac{(1-x)}{N-1}(V_{N-1} - V_1) + \delta x\frac{N-3}{N-1}(V_{N-3} - V_2)$$

The result follows, as we know from above that in the entrenched equilibrium, $u_i \leq 0$ for any $i \leq k-2$.

Any pure-strategy symmetric MPE in weakly undominated strategies belongs to exactly one of these two classes.

Let v (resp. w) denote the value brought to a member of the majority by the minority (resp. majority) candidate. So $v \in \{0, s\}$ and $w \in \{b, b+s\}$, and $v > w$ if and only if $(v, w) = (s, b)$ (otherwise $v < w$). Let $\mathcal{C} \equiv \left[0, ((\overline{x}+x)s + (1-x)b)/(1-\delta)\right]^k$. All vectors of value functions $(V_k, .., V_{N-1})$ necessarily belong to $\mathcal{C}$ as for any $s \geq b$, $\mathbb{E}_{v,w}[\max(v, w)] = (\overline{x}+x)s + (1-x)b$ [5]. By construction, given any $V_{k-1} \in \mathcal{C}$, the majority faces an optimal control problem, and there exists a unique sequence of majority value functions $(V_k(V_{k-1}), ..., V_{N-1}(V_{k-1}))$ solving the Bellman equations:

$$\forall i \geq k, \qquad V_i \equiv \mathbb{E}_{v,w}\Bigg[\max\Bigg\{v + \delta\Big(\frac{i-1}{N-1}V_{i-1} + \Big(1 - \frac{i-1}{N-1}\Big)V_i\Big), \\ w + \delta\Big(\frac{i}{N-1}V_i + \Big(1 - \frac{i}{N-1}\Big)V_{i+1}\Big)\Bigg\}\Bigg]$$

We note that, given any $V_{k-1} \in \mathcal{C}$, the majority can always guarantee a sequence of value functions such that $V_M \geq V_{k-1}$ for any $M \geq k$ (for instance by making meritocratic recruitments at all majority sizes, as such a strategy yields a flow payoff equal to $\mathbb{E}_{v,w}[\max(v, w)] \geq (1-\delta)V_{k-1}$). Hence in particular, any solution the Bellman equations given V_{k-1} satisfies $V_{k+1}(V_{k-1}) \geq V_{k-1}$, and therefore it is never optimal for a majority with size k to recruit the minority candidate whenever $v < w$.[6]

Hence, fix $V_{k-1} \in \mathcal{C}$ and consider the unique sequence of majority value functions $(V_M(V_{k-1}))_{M\geq k}$ solving the Bellman equations given V_{k-1}. If the following inequality holds:

$$s - b - \delta\frac{k-1}{N-1}(V_{k+1}(V_{k-1}) - V_{k-1}) < 0, \tag{8}$$

then any equilibrium yielding the sequence $(V_M(V_{k-1}))_{M\geq k}$ (if any) is k-entrenched. Moreover, by our initial remark, the majority candidate is recruited in $M = k$ whenever $v < w$, and thus $V_k(V_{k-1})$ writes as

$$V_k(V_{k-1}) = \mathbb{E}[w] + \delta\left[\frac{k}{N-1}V_k(V_{k-1}) + \frac{k-1}{N-1}V_{k+1}(V_{k-1})\right]$$

Consider the sequence of value functions $(V^{\mathrm{e}}_M(V_{k-1}))_{M\geq k}$ generated by the canonical entrenched strategy given V_{k-1}: the sequence $(V^{\mathrm{e}}_M(V_{k-1}))_{M\geq k}$ is defined recursively by the above

[5] Similarly, all vectors of minority value functions $(V_1, ..., V_{k-1})$ necessarily belong to $\mathcal{C}$.

[6] Indeed, the optimal recruitment for a majority with size k is determined by the sign of

$$v + \delta\left[\frac{k-1}{N-1}V_{k-1} + \frac{k}{N-1}V_k\right] - w - \delta\left[\frac{k}{N-1}V_k + \frac{k-1}{N-1}V_{k+1}\right] = v - w - \delta\frac{k-1}{N-1}(V_{k+1} - V_{k-1})$$

equation in $M = k$, while for any $M \geq k+1$,

$$V_M^{\mathrm{e}}(V_{k-1}) = \mathbb{E}_{v,w}[\max(v,w)] + \delta x \Big[\frac{M-1}{N-1} V_{M-1}^{\mathrm{e}}(V_{k-1}) + \frac{N-M}{N-1} V_M^{\mathrm{e}}(V_{k-1})\Big] + \delta(1-x)\Big[\frac{M}{N-1} V_M^{\mathrm{e}}(V_{k-1}) + \frac{N-M-1}{N-1} V_{M+1}^{\mathrm{e}}(V_{k-1})\Big]$$

The same computations[7] as in the proofs of Lemma 1 (see Online Appendix B) and Lemma C.1 (see Online Appendix C.2 below) then imply that the sequence of majority value functions $(V_M^{\mathrm{e}}(V_{k-1}))_{M \geq k}$ solves the Bellman equations given V_{k-1}, and therefore, as the latter have a unique solution, that $V_M(V_{k-1}) = V_M^{\mathrm{e}}(V_{k-1})$ for any $M \geq k$. The same computations then yield that for any $M \geq k+1$,[8]

$$w - v + \delta \Big[\frac{M-1}{N-1}(V_M(V_{k-1}) - V_{M-1}(V_{k-1})) + \frac{N-M-1}{N-1}(V_{M+1}(V_{k-1}) - V_M(V_{k-1}))\Big] \begin{cases} < 0 & \text{if } v > w \\ > 0 & \text{otherwise,} \end{cases}$$

and that whenever $v < w$,

$$w - v + \delta \Big[\frac{k-1}{N-1}(V_k(V_{k-1}) - V_{k-1}) + \frac{k-1}{N-1}(V_{k+1}(V_{k-1}) - V_k(V_{k-1}))\Big] > 0$$

Therefore, the canonical entrenchment strategies are strictly optimal at any majority size and for any realization of (v, w). The unique equilibrium such that (8) holds (if any) is thus the canonical entrenched equilibrium.

[7] Note that with such strategies, $V_M^{\mathrm{e}}(V_{k-1})$ does not depend on V_{k-1}, and thus $V_M^{\mathrm{e}}(V_{k-1}) = V_M^{\mathrm{e}}$ for any $M \geq k$. Hence, we may proceed as in the proof of Lemma C.1, using (5) to have that

$$\delta\left(\frac{k-2}{N-1} u_{k+1}^{\mathrm{e}} + \frac{k}{N-1} u_k^{\mathrm{e}}\right) = \frac{1}{1-x}\left(u_k^{\mathrm{e}} - x(s-b)\right)$$

and then the inequality $u_{k+1}^{\mathrm{e}} \leq u_k^{\mathrm{e}}$ from Lemma 1 in order to get that

$$\left[1 - \delta(1-x)\frac{N-2}{N-1}\right] u_k^{\mathrm{e}} \leq x(s-b)$$

and therefore, since $\delta < (N-1)/N$,

$$\delta\left(\frac{k-2}{N-1} u_{k+1}^{\mathrm{e}} + \frac{k}{N-1} u_k^{\mathrm{e}}\right) \leq \frac{\delta x \frac{N-2}{N-1}}{1 - \delta(1-x)\frac{N-2}{N-1}}(s-b) < s-b,$$

The result for any majority size $M \geq k+1$ follows by monotonicity of the sequence $(u_M^{\mathrm{e}})_M$ (established by Lemma 1).

[8] In particular, by Lemma 1, $V_M - V_{M-1} = V_M^{\mathrm{e}} - V_{M-1}^{\mathrm{e}} \geq 0$.

Similarly, and again fixing an arbitrary $V_{k-1} \in \mathcal{C}$, if the following inequality holds:

$$s - b - \delta \frac{k-1}{N-1}(V_{k+1}(V_{k-1}) - V_{k-1}) > 0, \tag{9}$$

then any equilibrium yielding the sequence $(V_M(V_{k-1})_{M \geq k}$ (if any) is k-meritocratic. By our initial remark, the majority candidate is recruited whenever $v < w$, and so any solution to the Bellman equations must satisfy:

$$V_k(V_{k-1}) = \mathbb{E}_{v,w}[\max(v,w)] + \delta x \left[\frac{k-1}{N-1} V_{k-1}(V_{k-1}) + \frac{k}{N-1} V_k(V_{k-1})\right] + \delta(1-x)\left[\frac{k}{N-1} V_k(V_{k-1}) + \frac{k-1}{N-1} V_{k+1}(V_{k-1})\right]$$

Consider the sequence of value functions $(V_M^{\mathrm{m}}(V_{k-1}))_{M \geq k}$ generated by the canonical meritocratic strategy given V_{k-1} [9]: the sequence $(V_M^{\mathrm{m}}(V_{k-1}))_{M \geq k}$ is defined recursively by

$$\forall M \geq k, \quad V_M^{\mathrm{m}}(V_{k-1}) = \mathbb{E}_{v,w}[\max(v,w)] + \delta x \left[\frac{M-1}{N-1} V_{M-1}^{\mathrm{m}}(V_{k-1}) + \frac{N-M}{N-1} V_M^{\mathrm{m}}(V_{k-1})\right] + \delta(1-x)\left[\frac{M}{N-1} V_M^{\mathrm{m}}(V_{k-1}) + \frac{N-M-1}{N-1} V_{M+1}^{\mathrm{m}}(V_{k-1})\right]$$

Using that as $V_{k-1} \in \mathcal{C}$, $V_k^{\mathrm{m}}(V_{k-1}) \geq V_{k-1}$, we have by the same computations as in the proof of Lemma 1 that:

$$0 \leq V_{N-1}^{\mathrm{m}}(V_{k-1}) - V_{N-2}^{\mathrm{m}}(V_{k-1}) \leq \ldots \leq V_k^{\mathrm{m}}(V_{k-1}) - V_{k-1}.$$

Moreover, since by assumption $V_{k+1}(V_{k-1})$ is the solution to the Bellman equations given V_{k-1}, we have that $V_{k+1}(V_{k-1}) \geq V_{k+1}^{\mathrm{m}}(V_{k-1})$, and thus (9) holds with $V_{k+1}^{\mathrm{m}}(V_{k-1})$. Consequently, using the monotonicity of the sequence $\left(V_M^{\mathrm{m}}(V_{k-1}) - V_{M-1}^{\mathrm{m}}(V_{k-1})\right)_{M \geq k}$, the argument of Lemma C.1 applies. Therefore, the sequence of value functions $(V_M^{\mathrm{m}}(V_{k-1}))_{M \geq k}$ generated by the canonical meritocratic strategy given V_{k-1} solves the Bellman equations, and as the latter have a unique solution, $V_i(V_{k-1}) = V_i^{\mathrm{m}}(V_{k-1})$ for any $i \geq k$. Using the same argument as above (relying on computations in Lemma C.1, together with Lemma 1), we can identify strategies and value functions, and thus the unique equilibrium such that (9) holds (if any) is the canonical meritocratic equilibrium.

Lastly, if neither (8) nor (9) hold, then

$$s - b - \delta \frac{k-1}{N-1}(V_{k+1}(V_{k-1}) - V_{k-1}) = 0,$$

[9] *In equilibrium*, $V_{k-1} = V_{k-1}^{\mathrm{m}}$, and thus $V_M^{\mathrm{m}}(V_{k-1}^{\mathrm{m}}) = V_M^{\mathrm{m}}$ with the previous notation.

i.e. in any corresponding equilibrium, the majority is indifferent between $\sigma(k) = 0$ and $\sigma(k) = 1$. As the Bellman equations have a unique solution, the above computations imply that the canonical entrenched and the canonical meritocratic strategies yield the same sequence of majority value functions $(V_M(V_{k-1}))_{M\geq k}$. Yet in equilibrium, the two strategies yield different values for V_{k-1}. Hence the above equality cannot hold in a pure-strategy equilibrium.

As a consequence, all k-entrenched (resp. k-meritocratic) equilibria are such that (8) (resp. (9)) holds. The above discussion thus implies that the canonical entrenched (resp. meritocratic) equilibrium is the unique k-entrenched (resp. k-meritocratic) equilibrium. Therefore, all symmetric Markov Perfect equilibria in weakly undominated strategies are canonical, as was to be shown.

C.2 Proof of Proposition 1-(ii)-(iii)-(iv)

We now characterize the existence regions of the canonical equilibria. We begin by establishing a necessary and sufficient condition for each canonical equilibrium to exist.

C.2.1 A necessary and sufficient condition for existence

Lemma C.1. *There exists no profitable one-shot deviation from a canonical strategy at any majority size and for any realization of the candidates' vertical types if and only if there exists no profitable deviation when $M = k$ and the minority candidate is strictly more talented.*

Proof. Clearly, we only need to show the "if" part. In both canonical equilibria, $u_i \equiv V_{i+1} - V_i$ is positive and decreasing for $i \geq k$ from Lemma 1. Hence, for any profile of candidates' "values" $(v, w) \in \{(0, b), (0, b + s), (s, b + s)\}$, this implies that for any $M \geq k$,

$$\delta\Big(1 - \frac{M}{N-1}\Big)\Big[V_{M+1} - V_M\Big] + \delta\frac{M-1}{N-1}\Big[V_M - V_{M-1}\Big] \geq 0 \geq v - w,$$

and thus by construction, the canonical strategies are optimal at any majority size whenever the majority candidate is at least as talented as the minority one: the majority then optimally selects its own candidate. Hence it remains to show that the canonical strategies are optimal at any majority size $M \geq k+1$ whenever the minority candidate is strictly more talented (i.e. equivalently when $v > w$). We first show the result for the entrenchment equilibrium, before showing the one for the meritocratic equilibrium.

(a) Suppose entrenchment is optimal when $M = k$, i.e. the majority is better off voting for an untalented majority candidate against a talented minority one. Then, letting V^{e} denote the value function in the entrenchment equilibrium[10] and using the notation $u_i^{\mathrm{e}} \equiv V_{i+1}^{\mathrm{e}} - V_i^{\mathrm{e}}$,

[10] Hence with V_{k-1}^{e} determined by the minority's entrenchment strategy

(2) implies that:

$$\delta \frac{k-1}{N-1} u_k^{\mathrm{e}} + \delta \frac{k-1}{N-1} u_{k-1}^{\mathrm{e}} \geq s - b$$

Similarly, it is optimal for the majority to recruit a talented minority candidate against an untalented majority one at majority size M if and only if:

$$\delta \left(1 - \frac{M}{N-1}\right) u_M^{\mathrm{e}} + \delta \frac{M-1}{N-1} u_{M-1}^{\mathrm{e}} \leq s - b,$$

Equation (5) implies that

$$\delta \left(\frac{k-2}{N-1} u_{k+1}^{\mathrm{e}} + \frac{k}{N-1} u_k^{\mathrm{e}} \right) = \frac{1}{1-x} \left(u_k^{\mathrm{e}} - x(s-b) \right)$$

Using the previous equation together with the inequality $u_{k+1}^{\mathrm{e}} \leq u_k^{\mathrm{e}}$ from Lemma 1 yields

$$\left[1 - \delta(1-x) \frac{N-2}{N-1} \right] u_k^{\mathrm{e}} \leq x(s-b)$$

Therefore, since $\delta < (N-1)/N$,

$$\delta \left(\frac{k-2}{N-1} u_{k+1}^{\mathrm{e}} + \frac{k}{N-1} u_k^{\mathrm{e}} \right) \leq \frac{\delta x \frac{N-2}{N-1}}{1 - \delta(1-x)\frac{N-2}{N-1}} (s-b) < s - b,$$

and hence it is indeed optimal for the majority to pick a talented minority candidate against an untalented majority one when $M = k+1$. The result extends to any majority size $M \geq k+1$ by monotonicity of the sequence $(u_i^{\mathrm{e}})_i$ which decreases with respect to i.

(b) Suppose meritocracy is optimal when $M = k$, i.e. the majority is better off voting for a talented minority candidate against an untalented majority one. Hence, with the usual notation, (2) implies:

$$\delta \frac{k-1}{N-1} u_k^{\mathrm{m}} + \delta \frac{k-1}{N-1} u_{k-1}^{\mathrm{m}} \leq s - b$$

Therefore, by Lemma 1, the monotonicity of the sequence $(u_i^{\mathrm{m}})_i$ which decreases with respect to i yields that for any $M \geq k+1$,

$$\delta \left(1 - \frac{M}{N-1}\right) u_M^{\mathrm{m}} + \delta \frac{M-1}{N-1} u_{M-1}^{\mathrm{m}} \leq s - b,$$

which concludes the proof of the Lemma. □

C.2.2 Existence regions

Transition probabilities depend on one's perspective: either "objective" (i.e. the one of an outsider), or "subjective" (i.e. the one of a majority or minority member). This observation motivates our introducing the following notation: For any given group, we refer to the transition probability, say from group sizes i to j, *from a group member's perspective* as the probability that the group's size goes from i to j *conditional on this group member being still a member next period.*

For regime $\mathrm{r} \in \{\mathrm{e}, \mathrm{m}\}$, let $p^{\mathrm{r}}_{i,j}$ be the transition probability from a majority member's perspective, i.e. the probability that the majority size moves from $i \geq k$ to $j \in \{i-1, i, i+1\}$[11] (note that $p^{\mathrm{r}}_{i,j} = 0$ if $|i-j| > 1$) from one period to another conditional on the majority member still being in the organization in the following period[12] (which has probability $(N-1)/N$). Then, for any $M > k$ and in the entrenched equilibrium ($\mathrm{r} = \mathrm{e}$):

$$\begin{cases} p^{\mathrm{e}}_{M,M+1} = (1-x)\left(1 - \dfrac{M+1}{N}\right)\dfrac{N}{N-1} = (1-x)\left(1 - \dfrac{M}{N-1}\right) \\ p^{\mathrm{e}}_{M,M} = \left[(1-x)\dfrac{M}{N} + x\left(1 - \dfrac{M}{N}\right)\right]\dfrac{N}{N-1} = (1-x)\dfrac{M}{N-1} + x\left(1 - \dfrac{M-1}{N-1}\right) \\ p^{\mathrm{e}}_{M,M-1} = x\dfrac{M-1}{N}\dfrac{N}{N-1} = x\dfrac{M-1}{N-1} \end{cases} \tag{10}$$

and

$$\begin{cases} p^{\mathrm{e}}_{k,k+1} = \left(1 - \dfrac{k+1}{N}\right)\dfrac{N}{N-1} = 1 - \dfrac{k}{N-1} \\ p^{\mathrm{e}}_{k,k} = \dfrac{k}{N}\dfrac{N}{N-1} = \dfrac{k}{N-1} \\ p^{\mathrm{e}}_{k,k-1} = 0 \end{cases} \tag{11}$$

For any $i, j \in \{1, ..., N-1\}$ and $t \in \mathbb{N}_+$, let $\pi^{\mathrm{e}}_{i,j}(t)$ be the t-period transition probability from i to j in the entrenched equilibrium from a majority member's perspective[13]. In other words, $\pi^{\mathrm{e}}_{i,j}(t)$ is the probability that starting from i, the majority size is equal to j after t periods conditional on the majority member still being in the organization[14]. Hence, for any

[11] If $j = k-1$, then the majority becomes the minority and the new majority is of size k.

[12] Consistently with our notation throughout the paper, the conditioning on the majority member still being in the organization in the following period makes the relevant discount factor be δ, i.e. the life-adjusted discount factor. We could have equivalently written the *unconditioned* transition probabilities, i.e. the probability of majority size going from i to j *and* the majority member still being in the organization in the following period, which would have led to using the pure-time discount δ_0.

[13] We focus on the entrenched equilibrium for the sake of exposition. Transition probabilities in the meritocratic regime have a more complex expression, with $\pi^{\mathrm{m}}_{i,j}(t)$ defined as the t-period transition probability from i to j *from the perspective of a member of the group initially of size* i.

[14] So in particular $\pi^{\mathrm{e}}_{i,i}(0) = 1$ and $\pi^{\mathrm{e}}_{i,j}(0) = 0$ for any $j \neq i$.

$i \in \{k, ..., N-1\}$ and $t \geq 0$,

$$\pi^{\mathrm{e}}_{i,M}(t+1) = p^{\mathrm{e}}_{M-1,M}\pi^{\mathrm{e}}_{i,M-1}(t) + p^{\mathrm{e}}_{M,M}\pi^{\mathrm{e}}_{i,M}(t) + p^{\mathrm{e}}_{M+1,M}\pi^{\mathrm{e}}_{i,M+1}(t)$$

We similarly explicit the transition probabilities from the perspective of a minority member. Let $\hat{p}^{\mathrm{e}}_{i,j}$ be the transition probability from a minority member's perspective, i.e. the probability that the majority size moves from $i \geq k$ to j from one period to another conditional on the minority member still being in the organization in the following period (which has probability $(N-1)/N$). Then, for any $M > k$ and in the entrenched equilibrium:

$$\begin{cases} \hat{p}^{\mathrm{e}}_{M,M+1} = (1-x)\Big(1 - \dfrac{M+2}{N}\Big)\dfrac{N}{N-1} = (1-x)\Big(1 - \dfrac{M+1}{N-1}\Big) \\ \hat{p}^{\mathrm{e}}_{M,M} = \Big[(1-x)\dfrac{M+1}{N} + x\Big(1 - \dfrac{M+1}{N}\Big)\Big]\dfrac{N}{N-1} = (1-x)\dfrac{M+1}{N-1} + x\Big(1 - \dfrac{M}{N-1}\Big) \\ \hat{p}^{\mathrm{e}}_{M,M-1} = x\dfrac{M}{N}\dfrac{N}{N-1} = x\dfrac{M}{N-1} \end{cases} \tag{12}$$

and

$$\begin{cases} \hat{p}^{\mathrm{e}}_{k,k+1} = \Big(1 - \dfrac{k+2}{N}\Big)\dfrac{N}{N-1} = 1 - \dfrac{k+1}{N-1} \\ \hat{p}^{\mathrm{e}}_{k,k} = \dfrac{k+1}{N}\dfrac{N}{N-1} = \dfrac{k+1}{N-1} \\ \hat{p}^{\mathrm{e}}_{k,k-1} = 0 \end{cases} \tag{13}$$

For any $i, j \in \{1, ..., N-1\}$, and $t \in \mathbb{N}_+$, let $\hat{\pi}^{\mathrm{e}}_{i,j}(t)$ be the t-period transition probability from i to j in the entrenchment equilibrium from a minority member's perspective. In other words, $\hat{\pi}^{\mathrm{e}}_{i,j}(t)$ is the probability that starting from i, the minority size is equal to j after t periods conditional on the minority member still being in the organization. Hence, for any $i \in \{k, ..., N-1\}$ and $t \geq 0$,

$$\hat{\pi}^{\mathrm{e}}_{i,M}(t+1) = \hat{p}^{\mathrm{e}}_{M-1,M}\hat{\pi}^{\mathrm{e}}_{i,M-1}(t) + \hat{p}^{\mathrm{e}}_{M,M}\hat{\pi}^{\mathrm{e}}_{i,M}(t) + \hat{p}^{\mathrm{e}}_{M+1,M}\hat{\pi}^{\mathrm{e}}_{i,M+1}(t)$$

For the meritocratic equilibrium, transition probabilities are given by (10) for majority members, and by (12) for minority members.

Note that because probabilities sum to 1,

$$\begin{cases} \left(\sum_{i=k+1}^{N-1} \hat{\pi}^{\mathrm{e}}_{k,i}(t)\right) - \left(\sum_{i=k+1}^{N-1} \pi^{\mathrm{e}}_{k+1,i}(t)\right) = -\left(\hat{\pi}^{\mathrm{e}}_{k,k}(t) - \pi^{\mathrm{e}}_{k+1,k}(t)\right) \\ \left(\sum_{i=k}^{N-1} \pi^{\mathrm{m}}_{k+1,i}(t)\right) - \left(\sum_{i=k}^{N-1} \pi^{\mathrm{m}}_{k-1,i}(t)\right) = -\left[\left(\sum_{i=1}^{k-1} \pi^{\mathrm{m}}_{k+1,i}(t)\right) - \left(\sum_{i=1}^{k-1} \pi^{\mathrm{m}}_{k-1,i}(t)\right)\right] \end{cases} \tag{14}$$

Proof of claims (ii) and (iii). We now turn to the statement of the existence result. Because the majority's choice when the majority candidate brings higher value than (or the same value as) the minority one is a no-brainer, let us examine the case in which the majority is tight and the minority candidate is more talented. The optimality decision hinges on the choice of the size and identity of the future majority.

Condition for existence of the meritocratic equilibrium. Leaving aside control considerations, choosing the less-deserving majority candidate when the majority is tight involves a cost $s - b$. To evaluate the impact of a potential switch of control, which will occur as we just saw with conditional probability $(k-1)/(N-1)$, note that in a meritocratic equilibrium, the present discounted expected quality of future appointees does not depend on the allocation of control. The only impact of the change in control is linked to homophily benefits when the two candidates have equal quality standing (which has probability $1 - 2x$), as control allows one to select the group's candidate. The present discounted probability of exercising control in future periods is higher if the majority keeps control next period than if it surrenders it. So overall a necessary condition of existence of a meritocratic equilibrium is:

$$s - b \geq \delta \frac{k-1}{N-1}(1-2x)b \sum_{t=0}^{+\infty} \delta^t \left[\left(\sum_{i=k}^{N-1} \pi^{\mathrm{m}}_{k+1,i}(t)\right) - \left(\sum_{i=k}^{N-1} \pi^{\mathrm{m}}_{k-1,i}(t)\right)\right]$$

And so the meritocratic equilibrium exists only if

$$\frac{s}{b} \geq \rho^{\mathrm{m}} \equiv 1 + \delta \frac{k-1}{N-1}(1-2x) \sum_{t=0}^{+\infty} \delta^t \left[\left(\sum_{i=k}^{N-1} \pi^{\mathrm{m}}_{k+1,i}(t)\right) - \left(\sum_{i=k}^{N-1} \pi^{\mathrm{m}}_{k-1,i}(t)\right)\right]$$

Lemma C.1 implies that this condition is in fact also sufficient: as is intuitive, deviations from meritocracy are less appealing further away from a tight majority size, i.e. from immediate control considerations.

Condition for existence of the entrenched equilibrium. Again, choosing the less talented majority candidate yields a direct payoff loss $s - b$. Then, with probability $(k-1)/(N-1)$, the surrendering of control translates into a permanent loss of homophily benefits whenever the two candidates have equal quality standing, which has probability $1 - 2x$. This cost is equal to

$$\frac{\delta}{1-\delta}(1-2x)b$$

Moreover, because the new majority will itself be entrenched, i.e. always voting for its own candidate whenever the majority is tight, the surrendering of control entails a further additional loss of homophily benefit proportional to $2xb$ whenever the majority is tight, along with the difference in homophily benefits associated with meritocratic decisions, i.e. choosing a talented minority candidate instead of an untalented majority candidate, at any majority size $M \geq k+1$. The latter would seem unwarranted as the two groups then agree on the decision to pick the more talented candidate; its existence comes from the fact that transition probabilities depend on one's perspective. Put together, these two terms add up to

$$\delta\frac{k-1}{N-1}2xb\sum_{t=0}^{+\infty}\delta^t\pi^{\mathrm{e}}_{k+1,k}(t) - \delta\frac{k-1}{N-1}xb\sum_{t=0}^{+\infty}\delta^t\left[\left(\sum_{i=k+1}^{N-1}\hat{\pi}^{\mathrm{e}}_{k,i}(t)\right) - \left(\sum_{i=k+1}^{N-1}\pi^{\mathrm{e}}_{k+1,i}(t)\right)\right]$$

Another way to interpret the homophily payoff terms consists in noticing that the expected per-period payoff of a majority (resp. minority) member is equal to $(1-x)b$ (resp. xb) whenever the majority is not tight ($M \geq k+1$), while it is equal to b (resp. 0) when majority is tight ($M = k$).

Finally, again because the new majority is itself entrenched, and since the shift in control implies that perspectives change, the surrendering of control yields a quality payoff equal to

$$\delta\frac{k-1}{N-1}(\overline{x}+x)s\sum_{t=0}^{+\infty}\delta^t\left[\left(\sum_{i=k+1}^{N-1}\hat{\pi}^{\mathrm{e}}_{k,i}(t)\right) - \left(\sum_{i=k+1}^{N-1}\pi^{\mathrm{e}}_{k+1,i}(t)\right)\right]$$
$$+\,\delta\frac{k-1}{N-1}\overline{x}s\sum_{t=0}^{+\infty}\delta^t\left(\hat{\pi}^{\mathrm{e}}_{k,k}(t) - \pi^{\mathrm{e}}_{k+1,k}(t)\right)$$

So overall a necessary condition for the existence of an entrenched equilibrium is

$$\begin{aligned}
b - s \geq\ & \delta\frac{k-1}{N-1}(\overline{x}+x)s\sum_{t=0}^{+\infty}\delta^t\left[\left(\sum_{i=k+1}^{N-1}\hat{\pi}^{\mathrm{e}}_{k,i}(t)\right) - \left(\sum_{i=k+1}^{N-1}\pi^{\mathrm{e}}_{k+1,i}(t)\right)\right] \\
& + \delta\frac{k-1}{N-1}\overline{x}s\sum_{t=0}^{+\infty}\delta^t\left(\hat{\pi}^{\mathrm{e}}_{k,k}(t) - \pi^{\mathrm{e}}_{k+1,k}(t)\right) - \frac{k-1}{N-1}\frac{\delta}{1-\delta}(1-2x)b \\
& - \delta\frac{k-1}{N-1}2xb\sum_{t=0}^{+\infty}\delta^t\pi^{\mathrm{e}}_{k+1,k}(t) + \delta\frac{k-1}{N-1}xb\sum_{t=0}^{+\infty}\delta^t\left[\left(\sum_{i=k+1}^{N-1}\hat{\pi}^{\mathrm{e}}_{k,i}(t)\right) - \left(\sum_{i=k+1}^{N-1}\pi^{\mathrm{e}}_{k+1,i}(t)\right)\right]
\end{aligned}$$

Let (15) be the inequality:

$$1+\delta\frac{k-1}{N-1}x\sum_{t=0}^{+\infty}\delta^t\Big(\pi^{\mathrm{e}}_{k+1,k}(t)-\hat{\pi}^{\mathrm{e}}_{k,k}(t)\Big)>0. \tag{15}$$

Define ρ^{e} as

$$\rho^{\mathrm{e}}\equiv\begin{cases}\dfrac{1+\dfrac{k-1}{N-1}\dfrac{\delta}{1-\delta}(1-2x)+\delta\dfrac{k-1}{N-1}x\sum_{t=0}^{+\infty}\delta^t\Big(\pi^{\mathrm{e}}_{k+1,k}(t)+\hat{\pi}^{\mathrm{e}}_{k,k}(t)\Big)}{1+\delta\dfrac{k-1}{N-1}x\sum_{t=0}^{+\infty}\delta^t\Big(\pi^{\mathrm{e}}_{k+1,k}(t)-\hat{\pi}^{\mathrm{e}}_{k,k}(t)\Big)} & \text{if (15) holds,}\\ +\infty \quad \text{otherwise.} & \end{cases}$$

Then the above argument suggests that the entrenched equilibrium exists only if $s/b\leq\rho^{\mathrm{e}}$. As the series term in (15) is negative for all t (see Lemma C.2 below), there might exist an entrenched equilibrium for all values of s and b (and in particular for $b=0$) for δ sufficiently close to 1, and thus we set $\rho^{\mathrm{e}}=+\infty$. Nonetheless, for a positive rate of time preference (which we assumed) – i.e. $\delta<(N-1)/N$ –, the entrenched equilibrium exists only on a finite interval: $\rho^{\mathrm{e}}<+\infty$.[15]

As hinted above, Lemma C.1 yields that these necessary conditions are also sufficient for these equilibria to exist. Hence, the entrenched (resp. meritocratic) equilibrium exists if and only if $s/b\leq\rho^{\mathrm{e}}$ (resp. $s/b\geq\rho^{\mathrm{m}}$).

Lastly, we show that the bounds ρ^{e} and ρ^{m} satisfy the following inequalities:[16]

$$1\leq 1+\delta\frac{k-1}{N-1}(1-2x)\leq\rho^{\mathrm{m}}\leq 1+\frac{\delta}{1-\delta}\frac{k-1}{N-1}(1-2x)<\rho^{\mathrm{e}}<+\infty \tag{16}$$

The upper and lower bounds on ρ^{m} may be decomposed as follows: $(1-2x)$ is the probability of a homophily benefit from control, $(k-1)/(N-1)$ the (conditional) probability of losing the majority when its end-of-period size is k, while δ (resp. $\delta/(1-\delta)$) are the time-discounted weights corresponding to a transient (resp. permanent) loss of control.[17]

The bounds on ρ^{e} and ρ^{m} in Inequality (16) derive from the following lemma.

Lemma C.2. *For all $t\geq 0$,*

(i) $\pi^{e}_{k+1,k}(t)\leq\hat{\pi}^{e}_{k,k}(t)$

[15] See Section C.3 for the proof of this result.

[16] The proof that $\rho^{\mathrm{e}}<+\infty$ is delayed to Section C.3.

[17] Note that ρ^{m} reaches its upper bound as x goes to 0. In the limit, it is equal to $1+\dfrac{\delta}{1-\delta}\dfrac{k-1}{N-1}$, which is intuitive: the majority weights the current-period payoff $s-b$ against the constant homophily loss in future periods due to the permanent loss of control (times its probability of occurrence $(k-1)/(N-1)$).

(ii) $\sum_{i \geq k} \pi^m_{k+1,i}(t) \geq \sum_{i \geq k} \pi^m_{k-1,i}(t)$

Proof. We use a result relying on the properties of monotone Markov chains.[18]

(i) Let P (resp. $\hat{P}$) be the stochastic matrix associated with the process $M(t)$ (resp. $\hat{M}(t)$) defined as the probability distribution over majority sizes $\{k, ..., N-1\}$ from a majority (resp. minority) member's perspective[19]. Namely, for any $i, j \in \{1, ..., k\}$,

$$P_{ij} = p^{\mathrm{e}}_{k+i-1,k+j-1}, \qquad \text{and} \qquad \hat{P}_{ij} = \hat{p}^{\mathrm{e}}_{k+i-1,k+j-1}$$

We first note that both P and $\hat{P}$ are *(strictly) stochastically-monotone* as $P_{i\cdot}$ stochastically dominates $P_{i'\cdot}$ whenever $i > i'$ (and similarly for $\hat{P}$)[20]. We then note that P and $\hat{P}$ are stochastically comparable, with $P_{i\cdot}$ stochastically dominating $\hat{P}_{i\cdot}$ for any $i \in \{1, ..., k\}$. Furthermore, the process $M(t)$ starts from the initial state $M(0) = (0, 1, 0, ...)$ which stochastically dominates the initial state of the process $\hat{M}(t)$, that is $\hat{M}(0) = (1, 0, , 0, ...)$.

Hence, a standard argument implies that for any $t > 0$, the distribution $M(t)$ stochastically dominates the distribution $\hat{M}(t)$ [21]. In particular, we have that for any $t > 0$,

$$\sum_{i=k+1}^{N-1} \pi^{\mathrm{e}}_{k+1,i}(t) \geq \sum_{i=k+1}^{N-1} \hat{\pi}^{\mathrm{e}}_{k,i}(t),$$

which, since probabilities sum to 1, is equivalent to: $\pi^{\mathrm{e}}_{k+1,k}(t) \leq \hat{\pi}^{\mathrm{e}}_{k,k}(t)$.

(ii) In order to establish the lower bound on ρ^{m} and thus Inequality (16), we note that:

$$\left(\sum_{i \geq k} \pi^{\mathrm{m}}_{k+1,i}(t)\right) - \left(\sum_{i \geq k} \pi^{\mathrm{m}}_{k-1,i}(t)\right) > 0 \qquad \forall t \geq 0$$

This inequality can be shown with the same technique as the one used in the proof of claim (i) by considering the process of one's successive in-group sizes in the meritocratic equilibrium, either starting from the initial state $k+1$ or $k-1$. Indeed, the same conditions are satisfied, as (a) both processes (of probability distribution over one's successive in-group sizes) share the

[18] See Daley, D. J. (1968). "Stochastically Monotone Markov Chains". *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 10(4):305-317; Keilson, J. and Kester, A. (1977). "Monotone Matrices and Monotone Markov Processes". *Stochastic Processes and their Applications*, 5(3):231-241.

[19] The i-th component of $M(t)$ is the probability (from the perspective of a majority member) that the majority be of size $k+1-i$ at period t. In particular, if at time 0 the majority is known to have size $k+1$, then $M(0) = (0, 1, 0, ..., 0)$, and at any later time t, $M(t) = \left(\pi^{\mathrm{e}}_{k+1,k}(t), ..., \pi^{\mathrm{e}}_{k+1,N-1}(t)\right)$. Similarly, if at time 0 the majority is known to have size k, then $\hat{M}(0) = (1, 0, , ..., 0)$, and at any later time t, $\hat{M}(t) = \left(\hat{\pi}^{\mathrm{e}}_{k,k}(t), ..., \hat{\pi}^{\mathrm{e}}_{k,N-1}(t)\right)$.

[20] Namely, for any $j^* \in \{1, ..., k\}$, $\Sigma_{j \geq j^*} P_{ij} \geq \Sigma_{j \geq j^*} P_{i'j}$.

[21] A sketch of the proof is as follows. Proceed by induction on t. The result for $t = 0$ holds as noted in the text. Suppose that $M(t)$ stochastically dominates $\hat{M}(t)$. Then, since P stochastically dominates $\hat{P}$, we have that $\hat{M}(t)P$ stochastically dominates $\hat{M}(t)\hat{P}$. Since P is stochastically-monotone, $M(t)P$ stochastically dominates $\hat{M}(t)P$. Thus, by transitivity, $M(t)P$ stochastically dominates $\hat{M}(t)\hat{P}$. In other words, $M(t+1)$ stochastically dominates $\hat{M}(t+1)$, which concludes the proof.

same transition matrix[22] which is stochastically monotone, and (b) the initial state with mass 1 in $k+1$ stochastically dominates the initial state with mass 1 in $k-1$. Hence the stochastic-comparison argument applies, yielding that the process of one's in-group size starting from $k+1$ stochastically dominates at any time $t \geq 0$ the process starting from $k-1$, and thus in particular,

$$\sum_{i \geq k} \pi^{\mathrm{m}}_{k+1,i}(t) > \sum_{i \geq k} \pi^{\mathrm{m}}_{k-1,i}(t)$$

□

Proof of claim (iii). The result derives from the explicit expressions of the existence thresholds together with Lemma C.2. Indeed, by the proof of Lemma C.2 and Proposition 1, we have that for all $t \geq 0$,

$$\pi^{\mathrm{e}}_{k+1,k}(t) - \hat{\pi}^{\mathrm{e}}_{k,k}(t) \leq 0, \qquad \text{and} \qquad \left(\sum_{i \geq k} \pi^{\mathrm{m}}_{k+1,i}(t)\right) - \left(\sum_{i \geq k} \pi^{\mathrm{m}}_{k-1,i}(t)\right) \geq 0$$

Using term-by-term differentiation of the series yields the result: $\partial \rho^{\mathrm{m}}/\partial \delta_0 \geq 0, \partial \rho^{\mathrm{e}}/\partial \delta_0 \geq 0$ for all $\delta_0 \in [0,1)$. Moreover, using term-by-term differentiation of the series for ρ^{m} and explicit computations for ρ^{e} yields

$$\left.\frac{\partial \rho^{\mathrm{m}}}{\partial \delta_0}\right|_{\delta_0=0} = \frac{k-1}{N}(1-2x) \qquad \text{and} \qquad \left.\frac{\partial \rho^{\mathrm{e}}}{\partial \delta_0}\right|_{\delta=0_0} = \frac{k-1}{N}$$

Lastly, the explicit expressions of the existence thresholds yield that for δ_0 close to 0, ρ^{m} and ρ^{e} increase with the size of the organization $N = 2k$.

C.3 Proof of Proposition 1: Entrenchment exists only on a finite interval ($\rho^{\mathrm{e}} < \infty$)

We show in this section that $\rho^{\mathrm{e}} < \infty$ [23]. The result is trivial for $k = 2$ (e.g. using the explicit expression of ρ^{e}). Let $k \geq 3$. The result is obvious for $x = 0$ (again by the explicit expression of ρ^{e}). We thus consider the case $x > 0$.

Let V^{e}_i denote the value function in the entrenched equilibrium, and define as before $u^{\mathrm{e}}_i \equiv V^{\mathrm{e}}_{i+1} - V^{\mathrm{e}}_i$. Fix $s > 0$. For any $i \in \{1, ..., N-2\}$, u^{e}_i is clearly continuous with respect to $b \in [0, +\infty)$.

The (one-shot) deviation differential payoff from entrenchment to meritocracy in $M = k$

[22] Namely, the matrix P^{m} with components $P_{ij} = p^{\mathrm{m}}_{i,j}$ for any $i, j \in \{1, ..., N-1\}$.

[23] The proof also yields that $\rho^{\mathrm{e}\dagger}|_{s^\dagger > b} < \infty$ (thus in particular for $x^\dagger \geq 1/2$).

writes:

$$s - b + \delta\frac{k-1}{N-1}(V_{k-1} - V_{k+1}) = s - b - \delta\frac{k-1}{N-1}(u^{\mathrm{e}}_{k-1} + u^{\mathrm{e}}_{k})$$

Fix $b = 0$. If the above payoff is strictly positive for $b = 0$, then by continuity, it must be so on a neighbourhood of 0. Hence there exists $\overline{\rho} > 0$ such that for any $s/b > \overline{\rho}$, there exists a strictly profitable deviation from entrenchment to meritocracy, which yields the result: $\rho^{\mathrm{e}} < \infty$. We thus show that for $b = 0$:

$$s - \delta\frac{k-1}{N-1}(u^{\mathrm{e}}_{k-1} + u^{\mathrm{e}}_{k}) > 0 \tag{17}$$

Using (2)-(3) and (4)-(6), the above inequality can be written as

$$\frac{\delta x\frac{k-1}{N-1}}{1 - \delta x - \delta(1-x)\left[\frac{k+1}{N-1} + \frac{k-2}{N-1}a_{k+1}\right]} \tag{18}$$
$$\times\left(1 - \frac{\delta x\frac{k}{N-1}}{1 - \delta(1-x)\left[\frac{k}{N-1} + \frac{k-2}{N-1}b_{k+1}\right]} - \frac{\delta x\frac{k-2}{N-1}}{1 - \delta(1-x)\left[\frac{k+1}{N-1} + \frac{k-3}{N-1}b_{k+2}\right]}\right) < 1$$

where the vectors $(a_{k+l})_{l=1}^{k-2}$, $(b_{k+l})_{l=1}^{k-2}$ are defined recursively by

$$\begin{cases} a_{k+l} = \dfrac{\delta x\frac{k+l}{N-1}}{1 - \delta(1-x)\left[\frac{k+l+1}{N-1} + \frac{k-l-2}{N-1}a_{k+l+1}\right] - \delta x\frac{k-l-1}{N-1}} \\[2ex] a_{N-2} = \dfrac{\delta x\frac{N-2}{N-1}}{1 - \delta(1-x) - \delta\frac{x}{N-1}} \end{cases}$$

and

$$\begin{cases} b_{k+l} = \dfrac{\delta x\frac{k+l-1}{N-1}}{1 - \delta(1-x)\left[\frac{k+l}{N-1} + \frac{k-l-2}{N-1}b_{k+l+1}\right] - \delta x\frac{k-l-1}{N-1}} \\[2ex] b_{N-2} = \dfrac{\delta x\frac{N-3}{N-1}}{1 - \delta(1-x)\frac{N-2}{N-1} - \delta\frac{x}{N-1}} \end{cases}$$

Indeed, computations using (2)-(3) and (4)-(6) for the entrenchment equilibrium, give that:

$$\begin{cases} \left[1-\delta(1-x)\dfrac{k+1}{N-1}-\delta x\right](V^{\mathrm{e}}_{k+1}-V^{\mathrm{e}}_{k-1}) \\ \qquad\qquad = xs+\delta(1-x)\dfrac{k-2}{N-1}(V^{\mathrm{e}}_{k+2}-V^{\mathrm{e}}_{k-2})-\delta x\dfrac{k}{N-1}u^{\mathrm{e}}_k+\delta x\dfrac{k-2}{N-1}u^{\mathrm{e}}_{k-2} \\ V^{\mathrm{e}}_{k+2}-V^{\mathrm{e}}_{k-2}=a_{k+1}(V^{\mathrm{e}}_{k+1}-V^{\mathrm{e}}_{k-1}) \\ u^{\mathrm{e}}_{k+1}=b_{k+1}u^{\mathrm{e}}_k \\ u^{\mathrm{e}}_{k-3}=b_{k+2}u^{\mathrm{e}}_{k-2} \end{cases}$$

and thus, by rearranging[24], (17) is equivalent to (18).

Remark. By construction, (a_{k+l}) and (b_{k+l}) are increasing with l, and for any l, $b_{k+l} < a_{k+l} < 1$. Moreover, for any l, a_{k+l} and b_{k+l} are increasing with respect to x and δ.[25]

We show that for any $x \in [0,1]$ and $\delta \in [0,(N-1)/N)$, inequality (18) is satisfied[26]. The above remark on the properties of $(a_{k+l})_{l=1}^{k-2}$, $(b_{k+l})_{l=1}^{k-2}$, yields that the term on the first line (resp. second line) is strictly increasing (resp. decreasing) with respect to x and δ. Moreover, by continuity, (18) is clearly satisfied for (x,δ) in a neighbourhood of $(0,\delta)$, $(x,0)$, and most interestingly of $(1,(N-1)/N)$.

Using the inequality $b_{k+1} < b_{k+2} < 1$, a sufficient condition for (18) to be satisfied is

$$\begin{aligned} &\delta x\frac{k-1}{N-1}\left(1-\delta x\frac{N-2}{N-1}-\delta(1-x)\left[\frac{k}{N-1}+\frac{k-2}{N-1}b_{k+1}\right]\right) \\ &\quad /\left[\left(1-\delta(1-x)\left[\frac{k}{N-1}+\frac{k-2}{N-1}b_{k+1}\right]\right)\left(1-\delta x-\delta(1-x)\left[\frac{k+1}{N-1}+\frac{k-2}{N-1}a_{k+1}\right]\right)\right] < 1 \end{aligned} \tag{19}$$

[24] Using in particular that (2)-(3) imply:

$$\begin{cases} u^{\mathrm{e}}_k = xs+\delta(1-x)\left[\left(1-\dfrac{k+1}{N-1}\right)u^{\mathrm{e}}_{k+1}+\dfrac{k}{N-1}u^{\mathrm{e}}_k\right] \\ u^{\mathrm{e}}_{k-2} = -xs+\delta(1-x)\left[\dfrac{k+1}{N-1}u^{\mathrm{e}}_{k-2}+\left(1-\dfrac{k+2}{N-1}\right)u^{\mathrm{e}}_{k-3}\right] \end{cases}$$

[25] These results can be shown by downward induction starting from $l = N-2$.

[26] The case $x \geq 1/2$ is equivalent to the homogamic-evaluation-capability setting with $x^\dagger \geq 1/2$. Indeed, the homogamic-evaluation-capability equivalent of (17) is:

$$\frac{\delta x\dfrac{k-1}{N-1}}{1-\delta x^\dagger-\delta(1-x^\dagger)\left[\dfrac{k+1}{N-1}+\dfrac{k-2}{N-1}a^\dagger_{k+1}\right]} \times\left(1-\frac{\delta x^\dagger\dfrac{k}{N-1}}{1-\delta(1-x^\dagger)\left[\dfrac{k}{N-1}+\dfrac{k-2}{N-1}b^\dagger_{k+1}\right]}-\frac{\delta x^\dagger\dfrac{k-2}{N-1}}{1-\delta(1-x^\dagger)\left[\dfrac{k+1}{N-1}+\dfrac{k-3}{N-1}b^\dagger_{k+2}\right]}\right) < \frac{x}{x^\dagger}$$

with the corresponding families $(a^\dagger_{k+l})_{l=1}^{k-2}$, $(b^\dagger_{k+l})_{l=1}^{k-2}$ defined as before by replacing x with $x^\dagger$.

or equivalently,

$$\delta x\frac{k-1}{N-1}\left(1-\delta x\frac{N-2}{N-1}-\delta(1-x)\left[\frac{k}{N-1}+\frac{k-2}{N-1}b_{k+1}\right]\right) \tag{20}$$
$$-\left(1-\delta(1-x)\left[\frac{k}{N-1}+\frac{k-2}{N-1}b_{k+1}\right]\right)\left(1-\delta x-\delta(1-x)\left[\frac{k+1}{N-1}+\frac{k-2}{N-1}a_{k+1}\right]\right)<0$$

The above inequalities are strictly stronger than (18) for any $x \in (0,1)$, and coincide with (18) in $x=1$. Moreover, it holds for any $(x,\delta)=(1,\delta)$ where $\delta \in [0,(N-1)/N]$.

We show that for any $x \in [0,1]$, (i) the LHS in (20) increases with δ over $[0,(N-1)/N]$, and (ii) this maximum (LHS with $\delta=(N-1)/N$) is strictly negative.

(i) In order to alleviate the notation, let C_a and C_b be defined as

$$C_a \equiv \frac{k+1}{N-1}+\frac{k-2}{N-1}a_{k+1}, \qquad \text{and} \qquad C_b \equiv \frac{k}{N-1}+\frac{k-2}{N-1}b_{k+1}$$

Since $b_{k+1}<a_{k+1}<1$, we have that $C_b<C_a<1$. Using a downward induction argument on the sequences $(a_{k+l})_l$ and $(b_{k+l})_l$ yields that $\partial a_{k+1}/\partial\delta > \partial b_{k+1}/\partial\delta$.[27] As a consequence,

$$\begin{aligned}\phi(\delta) &\equiv \frac{\partial a_{k+1}}{\partial\delta}\Big[1-\delta(1-x)C_b\Big]+\frac{\partial b_{k+1}}{\partial\delta}\left[1-\delta(1-x)C_a-\delta x\left(1+\frac{k-1}{N-1}\right)\right]\\ &\geq \frac{\partial b_{k+1}}{\partial\delta}\left[2-\delta(1-x)(C_a+C_b)-\delta x\left(1+\frac{k-1}{N-1}\right)\right]>0\end{aligned}$$

Denoting by $\varphi(\delta)$ the partial derivative of the LHS in (20) with respect to δ, we have after

[27] The result follows from the observation that

$$\frac{\partial a_{N-2}}{\partial\delta}=\frac{x\frac{N-2}{N-1}}{\left(1-\delta(1-x)-\delta\frac{x}{N-1}\right)^2}>\frac{x\frac{N-3}{N-1}}{\left(1-\delta(1-x)\frac{N-2}{N-1}-\delta\frac{x}{N-1}\right)^2}=\frac{\partial b_{N-2}}{\partial\delta}$$

and for any $l \in \{1,...,k-3\}$,

$$\begin{aligned}\frac{\partial a_{k+l}}{\partial\delta} &= \frac{x\frac{k+l}{N-1}+\delta^2x(1-x)\frac{k+l}{N-1}\frac{k-l-2}{N-1}\frac{\partial a_{k+l+1}}{\partial\delta}}{\left(1-\delta(1-x)\left[\frac{k+l+1}{N-1}+\frac{k-l-2}{N-1}a_{k+l+1}\right]-\delta x\frac{k-l-1}{N-1}\right)^2}\\ &> \frac{x\frac{k+l-1}{N-1}+\delta^2x(1-x)\frac{k+l-1}{N-1}\frac{k-l-2}{N-1}\frac{\partial b_{k+l+1}}{\partial\delta}}{\left(1-\delta(1-x)\left[\frac{k+l}{N-1}+\frac{k-l-2}{N-1}b_{k+l+1}\right]-\delta x\frac{k-l-1}{N-1}\right)^2}=\frac{\partial b_{k+l}}{\partial\delta}\end{aligned}$$

rearranging:

$$\varphi(\delta) = x\left(1+\frac{k-1}{N-1}\right) + (1-x)(C_a+C_b)$$
$$-2\delta\left[x(1-x)\left(1+\frac{k-1}{N-1}\right)C_b + (1-x)^2C_aC_b + x^2\frac{k-1}{N-1}\frac{N-2}{N-1}\right]$$
$$+\delta(1-x)\frac{k-2}{N-1}\left(\frac{\partial a_{k+1}}{\partial\delta}\Big[1-\delta(1-x)C_b\Big] + \frac{\partial b_{k+1}}{\partial\delta}\left[1-\delta(1-x)C_a - \delta x\left(1+\frac{k-1}{N-1}\right)\right]\right)$$

Let $\psi(\delta) \equiv \varphi(\delta) - \delta(1-x)\frac{k-2}{N-1}\phi(\delta)$. We then note that $\psi(\delta) \geq 0$ [28], and therefore, $\varphi(\delta) > 0$ for any $x \in [0,1]$. Consequently, the LHS in (20) is strictly increasing with respect to δ, and thus reaches its maximum over $[0,(N-1)/N]$ in $\delta = (N-1)/N$.

(ii) We now let $\delta = (N-1)/N$ and show that the LHS in (20) with $\delta = (N-1)/N$ is strictly negative. Indeed, the latter then writes as

$$LHS \equiv x\frac{k-1}{N}\left(1 - x\frac{N-2}{N} - (1-x)\left[\frac{k}{N}+\frac{k-2}{N}b_{k+1}\right]\right)$$
$$-\left(1-(1-x)\left[\frac{k}{N}+\frac{k-2}{N}b_{k+1}\right]\right)\left(1-x\frac{N-1}{N}-(1-x)\left[\frac{k+1}{N}+\frac{k-2}{N}a_{k+1}\right]\right)$$
$$= x\frac{k-1}{N}\left(\frac{1}{N}+(1-x)\frac{k-2}{N}(a_{k+1}-b_{k+1})\right)$$
$$-\left(\frac{k+1}{N}-\frac{1-x}{N}-(1-x)\frac{k-2}{N}b_{k+1}\right)\left(1-x\frac{N-1}{N}-(1-x)\left[\frac{k+1}{N}+\frac{k-2}{N}a_{k+1}\right]\right)$$

where b_{k+1} and a_{k+1} are evaluated in $\delta = (N-1)/N$. By using that $b_{k+1} < 1$ and rearranging,

[28] Indeed, the expressions of ϕ and φ yield after rearranging:

$$\psi(\delta) = x\left(1+\frac{k-1}{N-1}\right) + (1-x)(C_a+C_b) - 2\delta\left[x(1-x)\left(1+\frac{k-1}{N-1}\right)C_b + (1-x)^2C_aC_b + x^2\frac{k-1}{N-1}\frac{N-2}{N-1}\right]$$
$$= x\left[1+\frac{k-1}{N-1} - \delta(1-x)\left(1+\frac{k-1}{N-1}\right)C_b - \delta x\left(\frac{N-2}{N-1}\right)^2\right]$$
$$+(1-x)\left[\Big(C_a - \delta x C_b - \delta(1-x)C_aC_b\Big) + \left(C_b - \delta x\frac{k-1}{N-1}C_b - \delta(1-x)C_aC_b\right)\right] \geq 0$$

where the last inequality stems from the fact that $k/(N-1) < C_b < C_a < 1$.

we get that

$$
\begin{aligned}
LHS \leq\ & x\frac{k-1}{N}\left(\frac{1}{N}+(1-x)\frac{k-2}{N}(a_{k+1}-b_{k+1})\right) \\
& -\left(\frac{2}{N}+x\frac{k-1}{N}\right)\left(1-x\frac{N-1}{N}-(1-x)\left[\frac{k+1}{N}+\frac{k-2}{N}a_{k+1}\right]\right) \\
=\ & -\frac{2}{N^2}-(1-x)\frac{2}{N}\frac{k-2}{N}[1-a_{k+1}]-x(1-x)\frac{k-1}{N}\frac{k-2}{N}\Big[1-2a_{k+1}+b_{k+1}\Big]
\end{aligned}
$$

Hence, a sufficient condition for the LHS in (20) to be strictly negative is that $1-2a_{k+1}+b_{k+1}>0$. This actually holds, which concludes the proof: it can in fact be shown that for any $l\in\{1,...,k-2\}$, $1-2a_{k+l}+b_{k+l}\geq 0$ (with strict inequality whenever $x<1$).[29]

D Non-linear homophily benefits

A non-linear homophily benefit does not require enlarging the state space, as the size of the majority is still a sufficient statistics looking forward. While the homophily benefit of an extra in-group member depends on future hirings under a non-linear homophily benefit, the key trade-offs (driven by meritocracy vs. control) are not affected.

Let $\tilde{\mathcal{B}}(i)$ denote the per-period homophily benefit enjoyed by a member whose in-group has size i (thus, in the linear case, $\tilde{\mathcal{B}}(i)\equiv(i-1)\tilde{b}$).

[29] The argument is as follows. One first notes that since for any $l\in\{1,...,k-2\}$, $\partial a_{k+l}/\partial\delta\geq\partial b_{k+l}/\partial\delta>0$, the term $[1-2a_{k+l}+b_{k+l}]$ is strictly bounded below by its value for $\delta=(N-1)/N$. The rest of the argument derives from downward induction showing the result for any l with $\delta=(N-1)/N$. Explicit computations yield that for $\delta=(N-1)/N$,

$$
\big[1-2a_{N-2}+b_{N-2}\big]=\frac{(1-x)\frac{2}{N^2}}{\left(1-(1-x)\frac{N-1}{N}-\frac{x}{N}\right)\left(1-(1-x)\frac{N-2}{N}-\frac{x}{N}\right)}\geq 0
$$

Then, for any $l\in\{1,...,k-3\}$, the term $[1-2a_{k+l}+b_{k+l}]$ with $\delta=(N-1)/N$ has the same sign as

$$
\begin{aligned}
&\left(1-(1-x)\left[\frac{k+l+1}{N}+\frac{k-l-2}{N}a_{k+l+1}\right]-x\frac{k-l-1}{N}\right)\left(1-(1-x)\left[\frac{k+l}{N}+\frac{k-l-2}{N}b_{k+l+1}\right]-x\frac{k-l-1}{N}\right) \\
&\quad -2x\frac{k+l}{N}\left(1-(1-x)\left[\frac{k+l}{N}+\frac{k-l-2}{N}b_{k+l+1}\right]-x\frac{k-l-1}{N}\right) \\
&\quad +x\frac{k+l-1}{N}\left(1-(1-x)\left[\frac{k+l+1}{N}+\frac{k-l-2}{N}a_{k+l+1}\right]-x\frac{k-l-1}{N}\right) \\
&=(1-x)\left[\frac{k-l-1}{N}-\frac{k-l-2}{N}a_{k+l+1}\right]\left[\frac{k-l}{N}-x\frac{k-l-2}{N}-(1-x)\frac{k-l-2}{N}b_{k+l+1}\right]+x(1-x)\frac{k+l-1}{N}\frac{2}{N} \\
&\quad +x(1-x)\frac{k+l-1}{N}\frac{k-l-2}{N}\big[1-2a_{k+l+1}+b_{k+l+1}\big] \\
&\geq x(1-x)\frac{k+l-1}{N}\frac{k-l-2}{N}\big[1-2a_{k+l+1}+b_{k+l+1}\big]
\end{aligned}
$$

(a) Concave homophily benefit. Suppose, first, that $\tilde{\mathcal{B}}(i)$ is concave in the number of in-group members i, and that $\tilde{\mathcal{B}}(k+1)-\tilde{\mathcal{B}}(k) < \tilde{s}$ (otherwise super-entrenchment obtains[30]). The same analysis as in Section 2.1 shows that the equilibrium is either meritocratic or entrenched.

The concavity of $\tilde{\mathcal{B}}(i)$ may favor entrenchment or meritocracy. If $\tilde{\mathcal{B}}(i) \equiv (i-1)\tilde{b}$ for $i \geq k$, then concavity favors entrenchment, as the payoff under entrenchment is the same as in the fully-linear-homophily-benefit case, while the payoff under meritocracy is smaller. Symmetrically, if $\tilde{\mathcal{B}}(i) \equiv (i-1)\tilde{b}$ for $i \leq k$, the benefits of entrenchment are smaller under concavity while losing control has identical costs. Overall, concavity has ambiguous effects on the prevalent equilibrium.

(b) Convex homophily benefit. The analysis requires some adaptation in the case of convex homophily benefits[31]. We do not offer a full analysis, and content ourselves with the following observation (noted in the text). Suppose that the homophily benefit is linear up to $i = N-1$, but a large payoff accrues from full homogeneity (so that $\tilde{\mathcal{B}}(N) - \tilde{\mathcal{B}}(N-1)$ is larger than $\tilde{s}$). Then the organization may be meritocratic for small majorities and no longer so for large ones: the majority's expected cost of building a full majority (and maintaining it thereafter) becomes smaller as the majority size increases. While the ergodic state exhibits full entrenchment, the dynamics differ from the other instances of full entrenchment exhibited in the paper and may be meritocratic for a while.

E Proof of Proposition 2

We first show the result for majority members. For any $i \in \{k, ..., N-1\}$, let $v_i \equiv V_i^{\mathrm{m}} - V_i^{\mathrm{e}}$. By construction, for any $i \geq k+1$, the recursive expressions of V_i^{m} and V_i^{e} yield:

$$\left[1-\delta(1-x)\frac{i}{N-1} - \delta x\left(1-\frac{i-1}{N-1}\right)\right]v_i = \delta(1-x)\left(1-\frac{i}{N-1}\right)v_{i+1} + \delta x\frac{i-1}{N-1}v_{i-1}, \quad (21)$$

while for $i = k$,

$$v_k = \Delta + \delta\left[\frac{k}{N-1}v_k + \left(1-\frac{k}{N-1}\right)v_{k+1}\right]$$

where $\Delta \equiv x(s-b)+\delta x\frac{k-1}{N-1}(V_{k-1}^{\mathrm{m}} - V_{k+1}^{\mathrm{m}}) \geq 0$ – this last inequality stems from the definition of the meritocratic equilibrium (by the proof of Lemma 1, it is strict whenever $x > 0$ and either

[30] Namely, if there exists $l \in \{1, ..., k-1\}$ such that

$$\tilde{\mathcal{B}}(k+l) - \tilde{\mathcal{B}}(k+l-1) > \tilde{s} > \tilde{\mathcal{B}}(k+l+1) - \tilde{\mathcal{B}}(k+l),$$

then the only equilibrium is super-entrenchment at level l. Indeed, the myopically optimal choice allows the majority to keep control. Hence it can guarantee itself the upper bound on its payoff.

[31] Convexity may arise for instance when facilities or regulations must be added to accommodate the existence of a minority, or if a group's reaching a critical size delivers additional opportunities to its members.

$b > 0$ or $s > b$) –, and thus

$$\left[1 - \delta\frac{k}{N-1}\right]v_k = \Delta + \delta\Big(1 - \frac{k}{N-1}\Big)v_{k+1} \tag{22}$$

Assume by contradiction that $v_{N-1} < 0$. Then, Equation (21) for $i = N-1$ implies that $v_{N-2} < v_{N-1} < 0$, and thus by induction that $v_k < v_{k+1} < ... < v_{N-1} < 0$. However, Equation (22) then yields $0 > (1-\delta)v_k > \Delta \geq 0$, which is a contradiction. Hence, $v_{N-1} \geq 0$, and by induction using Equation (21), $v_k \geq v_{k+1} \geq ... \geq v_{N-1} \geq 0$, which concludes the proof. Note that the inequalities are strict whenever $\Delta > 0$, i.e. whenever $b > 0$ and $x > 0$ (with $s > b$ if $x = 1/2$).

The result for minority members follows from analogous computations, noting that for $M \geq k+1$, meritocracy and entrenchment yield the same flow payoffs and transition probabilities, while in $M = k$,

$$\begin{aligned} V^{\mathrm{m}}_{k-1} - V^{\mathrm{e}}_{k-1} = x(s+b) + x\delta\Big[\frac{k-2}{N-1}(V^{\mathrm{m}}_{k-1} - V^{\mathrm{m}}_{k-2}) + \frac{k}{N-1}(V^{\mathrm{m}}_{k} - V^{\mathrm{m}}_{k-1})\Big] \\ + \delta\Big[\frac{k-2}{N-1}(V^{\mathrm{m}}_{k-2} - V^{2}_{k-2}) + \frac{k+1}{N-1}(V^{\mathrm{m}}_{k-1} - V^{\mathrm{e}}_{k-1})\Big], \end{aligned}$$

where $V^{\mathrm{m}}_{k-2} \leq V^{\mathrm{m}}_{k-1} \leq V^{\mathrm{m}}_{k}$ by Lemma 1. Hence, $V^{\mathrm{m}}_i \geq V^{\mathrm{e}}_i$ for any $i \leq k-1$.

Lastly, as a by-product of the proof, we have that the gap between the value functions in the two equilibria, $V^{\mathrm{m}}_i - V^{\mathrm{e}}_i$, decreases as the majority size moves further away from $M = k$.[32]

F Proof of Lemma 2

We show successively that:

(i) $\nu^{\mathrm{e}}_k = 0$

(ii) for any $i \geq k+1$, we have that: $\dfrac{\nu^{\mathrm{e}}_{i+1}}{\nu^{\mathrm{e}}_i} = \dfrac{\nu^{\mathrm{m}}_{i+1}}{\nu^{\mathrm{m}}_i} = \dfrac{1-x}{x}\dfrac{N-i}{i+1}$,

(iii) $\nu^{\mathrm{e}}_k + \nu^{\mathrm{e}}_{k+1} < \nu^{\mathrm{m}}_k + \nu^{\mathrm{m}}_{k+1}$

and so, that the probability distribution $\{\nu^{\mathrm{e}}_i\}$ strictly first-order stochastically dominates $\{\nu^{\mathrm{m}}_i\}$.

Claim *(i)* derives from the fact that i refers to the size of the majority at the end of the period $i \in \{k, ..., 2k\}$. Note that in regime $\mathrm{r} \in \{\mathrm{e}, \mathrm{m}\}$,

$$\nu^{\mathrm{r}}_N = (1-x)\nu^{\mathrm{r}}_N + \frac{1-x}{N}\nu^{\mathrm{r}}_{N-1}$$

and for $k+2 \leq i < N$, $\quad \nu^{\mathrm{r}}_i = (1-x)\dfrac{N-(i-1)}{N}\nu^{\mathrm{r}}_{i-1} + \Big[(1-x)\dfrac{i}{N} + x\dfrac{N-i}{N}\Big]\nu^{\mathrm{r}}_i + x\dfrac{i+1}{N}\nu^{\mathrm{r}}_{i+1}$

[32] The result for $i \leq k-1$ can be established using analogous computations to the case $i \geq k$, relying on the recursive expressions of the minority value functions.

Claim *(ii)* follows by backward induction starting from $i = N$ and going down until $k+2$ included. Note that the explicit expression of the ergodic distribution in the entrenched equilibrium obtains with claims *(i)* and *(ii)* by writing $\Sigma_{i=k+1}^{N}\nu_i^{\mathrm{e}} = 1$. The explicit expression of the ergodic distribution in the meritocratic equilibrium obtains similarly noting that $(1-x)N\nu_k^{\mathrm{m}} = x(k+1)\nu_{k+1}^{\mathrm{m}}$. One has in particular that

$$\begin{cases} \nu_{k+1}^{\mathrm{m}}\Big[\dfrac{x}{1-x}\dfrac{k+1}{N} + 1 + \displaystyle\sum_{i=1}^{k-1}\Big(\dfrac{1-x}{x}\Big)^i \prod_{l=1}^{i}\dfrac{k-l}{k+1+l}\Big] = 1 \\ \nu_{k+1}^{\mathrm{e}}\Big[1 + \displaystyle\sum_{i=1}^{k-1}\Big(\dfrac{1-x}{x}\Big)^i \prod_{l=1}^{i}\dfrac{k-l}{k+1+l}\Big] = 1 \end{cases}$$

Lastly, claims *(i)* and *(ii)* together imply claim *(iii)*.

Remark. The ergodic probability for the majority size to be equal to k at the beginning of a period in the entrenched equilibrium writes as $\nu_{k+1}^{\mathrm{e}}(k+1)/N$, and thus by the above expression, decreases with k.

G Proof of Proposition 3

Let ρ^W be uniquely defined by

$$qN(N-1)\Big[1 + \frac{x}{1-x}\frac{k+1}{N} + \sum_{l=1}^{k-1}\Big(\frac{1-x}{x}\Big)^l \prod_{j=1}^{l}\frac{k-j}{k+1+j}\Big]\rho^W = \frac{2}{1-x}\Big[1 + \sum_{l=1}^{k-1}(l+1)^2\Big(\frac{1-x}{x}\Big)^l \prod_{j=1}^{l}\frac{k-j}{k+1+j}\Big]$$

We show that $W^{\mathrm{m}} \geq W^{\mathrm{e}}$ if and only if $s/b \geq \rho^W$. The result thus obtains by showing that $\rho^W < 1$ for all parameter values.

We first establish the explicit expression of ρ^W. By construction, we have that

$$B^{\mathrm{m}} - B^{\mathrm{e}} = \sum_{i=k}^{N}(\nu_i^{\mathrm{m}} - \nu_i^{\mathrm{e}})\Big[i(i-1) + (N-i)(N-i-1)\Big]\tilde{b}$$

Hence, computations using the explicit expressions of the ergodic distributions (see Section F above) yield after rearranging:

$$\Big[\frac{x}{1-x}\frac{k+1}{N} + 1 + \sum_{i=1}^{k-1}\Big(\frac{1-x}{x}\Big)^i \prod_{l=1}^{i}\frac{k-l}{k+1+l}\Big]\Big[1 + \sum_{i=1}^{k-1}\Big(\frac{1-x}{x}\Big)^i \prod_{l=1}^{i}\frac{k-l}{k+1+l}\Big](B^{\mathrm{m}} - B^{\mathrm{e}}) = -\frac{2x}{1-x}\frac{k+1}{N}\Big[1 + \sum_{l=1}^{k-1}(l+1)^2\Big(\frac{1-x}{x}\Big)^l \prod_{j=1}^{l}\frac{k-j}{k+1+j}\Big]\tilde{b}$$

Similar computations for $(S^{\mathrm{m}} - S^{\mathrm{e}})$ yield:

$$\left[\frac{x}{1-x}\frac{k+1}{N} + 1 + \sum_{i=1}^{k-1}\left(\frac{1-x}{x}\right)^i \prod_{l=1}^{i} \frac{k-l}{k+1+l}\right]\left[1 + \sum_{i=1}^{k-1}\left(\frac{1-x}{x}\right)^i \prod_{l=1}^{i} \frac{k-l}{k+1+l}\right](S^{\mathrm{m}} - S^{\mathrm{e}})$$
$$= N(N-1)x\frac{k+1}{N}\left[\frac{x}{1-x}\frac{k+1}{N} + 1 + \sum_{i=1}^{k-1}\left(\frac{1-x}{x}\right)^i \prod_{l=1}^{i} \frac{k-l}{k+1+l}\right]\tilde{s}$$

The expression of ρ^W follows. Lastly, the inequality $\rho^W < 1$ derives from the observations that for any $x \in [0, 1/2]$, $N(N-1) > 2(l+1)^2/(1-x)$ for any $l \leq k-2$, and that[33]

$$N(N-1)\left[1 + \left(\frac{1-x}{x}\right)^{k-1} \prod_{l=1}^{k-1} \frac{k-l}{k+1+l}\right] > \frac{2}{1-x}\left[1 + k^2\left(\frac{1-x}{x}\right)^{k-1} \prod_{l=1}^{k-1} \frac{k-l}{k+1+l}\right]$$

H Proof of Proposition 4

We first show the validity of the remark in the text on a blind principal ($\lambda = 0$), before establishing Proposition 4.

If the principal does not observe horizontal types and in particular the majority size, it worsens the efficiency of its interventions as it cannot fine-tune its interventions. Hence if it is an equilibrium for the principal not to intervene when it observes horizontal types, it is also an equilibrium to do so when the principal is totally blind. We thus consider the case where the principal observes horizontal types and show that, taking as given members' beliefs on the principal's strategy, it is optimal for the latter not to engage in interventions. There is clearly no benefit for the principal to intervene whenever the majority is not tight ($M \geq k+1$) – or whenever it is tight and meritocratic – as then the majority's choice maximizes the organization's quality and, by resolving ties in favor of the majority candidate, also maximizes the homophily payoff conditional on maximizing the organization's quality. Hence, for $s > b$ and $q \geq 1$, the majority's choice is optimal from the principal's point of view.[34]

Thus we now need to show that it is optimal[35] for the principal not to intervene in the entrenchment equilibrium when majority is tight ($M = k$). Since a tight entrenched majority

[33] Indeed, as the inequality $N(N-1) < 2k^2/(1-x)$ holds if and only if $x > (k-1)/(N-1)$, we have that for any $x \in [0, 1/2]$, the difference between the LHS minus the RHS is bounded below by

$$N(N-1)\left[1 + \left(\frac{k}{k-1}\right)^{k-1} \prod_{l=1}^{k-1} \frac{k-l}{k+1+l}\right] - 4\left[1 + k^2\left(\frac{k}{k-1}\right)^{k-1} \prod_{l=1}^{k-1} \frac{k-l}{k+1+l}\right] > N(N-1) - 4 - N > 0$$

where the first inequality derives from $\left(\frac{k}{k-1}\right)^{k-1} \prod_{l=1}^{k-1} \frac{k-l}{k+1+l} < 1$, while the second holds for any $N \geq 4$.

[34] Fix $s > b$. Since the quality payoff accrues to all members of the organization, while the homophily benefit only accrues to the in-group members, this optimality persists for q in a lower neighbourhood of 1. Furthermore, the neighbourhood expands toward 0 as the ratio s/b increases.

[35] Strictly so if there is any small cost of intervention, or if the principal internalizes members' homophily benefits.

always votes for its own candidate, its vote carries no information on the candidates' respective talents: the principal cannot do better by observing horizontal types than it can without. Hence, from the quality perspective, the principal picks the (of "a" if there is a tie) right candidate with probability 1/2, whereas the majority does so with probability $(1-x) \geq 1/2$. Similarly, the majority takes the homophily-maximizing decision with probability 1, while the principal can at best replicate this probability if it observes horizontal types, and can only do so with probability 1/2 if it does not. Hence the principal cannot outperform the majority's decision.

We now turn to the proof of Proposition 4.

Proof of claim (i). Let $\lambda > 0$ be the probability that the principal learns the quality of the candidates. The proof unfolds in two steps:

(a) We show that for s/b sufficiently close to 1, there exists a profitable deviation from canonical entrenchment in $k+1$ (the unique outcome when s/b is close to 1 and $\lambda = 0$) toward super-entrenchment at level 1. The argument then extends to full-entrenchment.

(b) We show that for s/b sufficiently close to 1, there can be no profitable deviation from full entrenchment.

(a). For $i \geq k$, let V_i be the majority value function in the canonical entrenchment equilibrium with probability of intervention $\eta = x\lambda > 0$. In order to alleviate the notation, we drop the superscript e and the notation for the dependence on λ. Consider a deviation from canonical entrenchment to super-entrenchment in $k+1$, i.e. the majority voting its own, less talented candidate against the strictly more talented minority one, and being overruled with probability λ. The (one-shot) differential payoff from the deviation at $M = k+1$ writes

$$\begin{aligned}\Delta &\equiv (1-\lambda)\Big[b - s + \delta\Big(\frac{k+1}{N-1}V_{k+1} + \frac{k-2}{N-1}V_{k+2}\Big) - \delta\Big(\frac{k}{N-1}V_k + \frac{k-1}{N-1}V_{k+1}\Big)\Big] \\ &= (1-\lambda)\Big[b - s + \delta\Big(\frac{k-2}{N-1}u_{k+1} + \frac{k}{N-1}u_k\Big)\Big]\end{aligned}$$

where $u_i = V_{i+1} - V_i$. The sequence $(u_i)_{1\leq i\leq N-2}$ satisfies Equation (4) for any $i \geq k+1$, and Equation (6) for any $i \leq k-3$, while

$$\begin{cases}\Big[1 - \delta(1-x)\dfrac{k}{N-1} - \delta x\lambda\dfrac{k-1}{N-1}\Big]u_k = x(1-\lambda)(s-b) + \delta(1-x)\dfrac{k-2}{N-1}u_{k+1} + \delta x\lambda\dfrac{k-1}{N-1}u_{k-1} \\ \Big[1 - \delta(1-x\lambda)\Big]u_{k-1} = (1-2x\lambda)b + \delta(1-x\lambda)\Big[\dfrac{k-2}{N-1}u_{k-2} + \dfrac{k-1}{N-1}u_k\Big] \\ \Big[1 - \delta(1-x)\dfrac{k+1}{N-1} - \delta x\lambda\dfrac{k-2}{N-1}\Big]u_{k-2} = -x(1-\lambda)(s+b) + \delta(1-x)\dfrac{k-3}{N-1}u_{k-3} + \delta x\lambda\dfrac{k}{N-1}u_{k-1}\end{cases} \tag{23}$$

Summing up on all indices yields[36]

$$\left[1-\delta\frac{x}{N-1}-\delta(1-x)\right](u_1+u_{N-2})+(1-\delta)\sum_{i=2}^{N-3}u_i=(1-2x)b>0 \tag{24}$$

Fix $b>0$. For any $s\geq b$, the same argument as the one used in the proof of Lemma 1 yields $u_k>u_{k+1}>...>u_{N-2}>0$. Put succinctly, one supposes by contradiction that $u_{N-2}\leq 0$ and reaches a contradiction showing by induction, using (4) together with the above system, that this implies $u_{k-1}\leq 0$. Then, if $u_1\leq 0$, (6) implies $u_i\leq 0$ for all i, which contradicts (24); whereas if $u_1>0$, (6) implies $u_{k-1}>0$ and we reach again a contradiction. Hence $u_{N-2}>0$ and the same induction argument using (4) thus brings the result.

The differential deviation payoff is thus strictly positive if and only if

$$\delta\left(\frac{k-2}{N-1}u_{k+1}+\frac{k}{N-1}u_k\right)>s-b \tag{25}$$

Consequently, for $s=b$, (25) is satisfied as it writes

$$\delta\left(\frac{k-2}{N-1}u_{k+1}+\frac{k}{N-1}u_k\right)>0$$

Lastly, since for fixed b, $(u_i)_i$ is continuous with respect to s, this implies that for any s/b sufficiently close to 1, there exists a strictly profitable (one-shot) deviation from canonical entrenchment to super-entrenchment.

As a by-product of the proof, we have by the same argument that whenever $\eta=0$, there exists no profitable deviation from canonical entrenchment to super-entrenchment as then[37]

$$\delta\left[\frac{k-2}{N-1}u_{k+1}|_{\lambda=0}+\frac{k}{N-1}u_k|_{\lambda=0}\right]<\frac{x}{1-x}\left[\left(1-\delta(1-x)\frac{k}{N-1}\right)^{-1}-1\right](s-b)<s-b$$

The same argument shows that, for s/b sufficiently close to 1, there exist profitable deviations from any level $l\geq 0$ of entrenchment toward entrenchment at a higher level, and thus in particular toward full-entrenchment.

Lastly, we show that for s/b in a neighbourhood of 1, full-entrenchment equilibrium is the unique symmetric MPE in weakly undominated strategies. To this end, we show that, for s/b in a neighbourhood of 1, any symmetric MPE in weakly undominated strategies is monotonic, in the sense that a stronger majority makes more meritocratic recruitments. The result crucially

[36] Assuming $k\geq 4$. The expression for $k\in\{2,3\}$ writes differently on the LHS but has the same implication.
[37] Indeed,

$$\left[1-\delta(1-x)\frac{k}{N-1}\right]u_k\Big|_{\lambda=0}=x(s-b)+\delta(1-x)\left(1-\frac{k+1}{N-1}\right)u_{k+1}\Big|_{\lambda=0}<x(s-b)$$

relies on the fact that the minority is never pivotal – as opposed to the absenteeism (Section 4.2.2) setting.

Let $s = b > 0$. We show that in any symmetric MPE in weakly undominated strategies, the sequence of differential value function $(u_M)_{M\geq k-1}$ is strictly positive and strictly decreases with M. As a consequence, by continuity, it is so for s/b in a neighbourhood of 1, and this in turn implies that, for s/b in such a neighbourhood, any symmetric MPE in weakly undominated strategies is monotonic as differential payoffs in the trade-off between a meritocratic or an entrenched recruitment at majority size M (whenever the minority candidate is strictly more talented than the majority one) write as

$$s - b - \delta\left[\frac{M-1}{N-1}u_{M-1} + \left(1 - \frac{M}{N-1}\right)u_M\right]$$

and are thus increasing with M.

For $s = b > 0$, we have that

$$\begin{cases} u_{k-1} = (1-2x\lambda)b + \delta(1-x\lambda)\left[\frac{k-2}{N-1}u_{k-2} + u_{k-1} + \frac{k-1}{N-1}u_k\right] & \text{in an equilibrium in which the majority is entrenched in } k, \\ u_{k-1} = (1-2x)b + \delta(1-x)\left[\frac{k-2}{N-1}u_{k-2} + u_{k-1} + \frac{k-1}{N-1}u_k\right] & \text{in an equilibrium in which it is meritocratic in } k. \end{cases}$$

and for any majority size $M \leq N-2$,

$$\begin{cases} u_M = \delta(1-x\lambda)\left[\frac{M}{N-1}u_M + \left(1-\frac{M+1}{N-1}\right)u_{M+1}\right] + \delta x\lambda\left[\frac{M-1}{N-1}u_{M-1} + \left(1-\frac{M}{N-1}\right)u_M\right] \\ \qquad \text{in an equilibrium in which the majority is entrenched in } M, M+1, \\ u_M = \delta(1-x)\left[\frac{M}{N-1}u_M + \left(1-\frac{M+1}{N-1}\right)u_{M+1}\right] + \delta x\left[\frac{M-1}{N-1}u_{M-1} + \left(1-\frac{M}{N-1}\right)u_M\right] \\ \qquad \text{in an equilibrium in which the majority is meritocratic in } M, M+1, \\ u_M = \delta(1-x)\left[\frac{M}{N-1}u_M + \left(1-\frac{M+1}{N-1}\right)u_{M+1}\right] + \delta x\lambda\left[\frac{M-1}{N-1}u_{M-1} + \left(1-\frac{M}{N-1}\right)u_M\right] \\ \qquad \text{in an equilibrium in which the majority is entrenched (resp. meritocratic) in } M \text{(resp. } M+1), \\ u_M = \delta(1-x\lambda)\left[\frac{M}{N-1}u_M + \left(1-\frac{M+1}{N-1}\right)u_{M+1}\right] + \delta x\left[\frac{M-1}{N-1}u_{M-1} + u_M + \left(1-\frac{M+1}{N-1}\right)u_{M+1}\right] \\ \qquad \text{in an equilibrium in which the majority is meritocratic (resp. entrenched) in } M \text{(resp. } M+1), \end{cases}$$

together with similar expressions for u_i when $i \leq k-2$.

Hence, we apply the usual argument: supposing by contradiction that $u_{N-2} \leq 0$, and working by induction using the sums of u_i over appropriate indices in order to reach the contradiction – which ultimately derives from the fact that there is a unique flow differential payoff, and that it is equal either to $(1-2x\lambda)b$ or $(1-2x)b$, which are both strictly positive. This yields that in any equilibrium, $u_{N-2}>0$, and the above system then implies that $u_i>0$ for all $i\in\{k-1,...,N-2\}$ as was to be shown.

Non-ergodic welfare comparison and equilibrium selection. Proposition 2 yields that, whenever meritocracy co-exists with entrenchment, the former is preferred by all members of the organization at any majority size. The result goes through in this setting.

We now show that for any $l\geq 2$, whenever super-entrenchment at level $l-1$ and super-entrenchment at level l co-exist in equilibrium, the former is preferred by all (current) members of the organization at any majority size. The result for majority members relies on the same computations as in the proof of Proposition 2 (see Online Appendix E), using that since super-entrenchment at level $l-1$ is an equilibrium[38],

$$s-b+\delta\Big(\frac{k+l}{N-1}u^{\mathrm{e},l-1}_{k+l}+\frac{k-l-2}{N-1}u^{\mathrm{e},l-1}_{k+l+1}\Big)\geq 0$$

where $u^{\mathrm{e},l-1}_i=V^{\mathrm{e},l-1}_{i+1}-V^{\mathrm{e},l-1}_i$ with $V^{\mathrm{e},l-1}_i$ the value function of being in a group of size i in the super-entrenchment at level $l-1$ equilibrium. The result for minority members also relies on analogous computations to the ones in the proof of Proposition 2 (see Online Appendix E): using the recursive expressions of the value function for minority members in a similar fashion, we have that $V^{\mathrm{e},l-1}_i\geq V^{\mathrm{e},l}_i$ for any $i\leq k-1$ if

$$s+b+\delta\Big(\frac{k-l-1}{N-1}u^{\mathrm{e},l-1}_{k-l-1}+\frac{k+l-1}{N-1}u^{\mathrm{e},l-1}_{k-l}\Big)\geq 0 \tag{26}$$

We thus show that this inequality holds, using the recursive expressions of $(u^{\mathrm{e},l-1}_i)_i$. We distinguish two cases.
(1) if $u^{\mathrm{e},l-1}_{k-l}\geq 0$, then $0\geq u^{\mathrm{e},l-1}_1\geq u^{\mathrm{e},l-1}_2\geq ...\geq u^{\mathrm{e},l-1}_{k-l}$.[39] Hence inequality (26) holds.
(2) if $u^{\mathrm{e},l-1}_{k-l}\leq 0$, then $u^{\mathrm{e},l-1}_{k-l}\leq u^{\mathrm{e},l-1}_{k-l-1}$ and $u^{\mathrm{e},l-1}_{k-l}\leq u^{\mathrm{e},l-1}_{k-l+1}$. Indeed,

- consider the first inequality and suppose by contradiction that $u^{\mathrm{e},l-1}_{k-l-1}<u^{\mathrm{e},l-1}_{k-l}$. By the now-usual (contradiction and induction) argument, this implies that $u_1<...<u^{\mathrm{e},l-1}_{k-l}\leq 0$. However, by summing the recursive expressions of $u^{\mathrm{e},l-1}_i$ for $i=1,...k-l-1$, and

[38]Indeed, this implies that in equilibrium, meritocratic recruitments are the majority's best response whenever it has size $k+l$, hence the inequality.

[39]This can be shown by the now-usual argument, supposing by contradiction that $u^{\mathrm{e},l-1}_1<0$, which implies by the recursive expressions of $(u^{\mathrm{e},l-1}_i)_i$, that $0>u^{\mathrm{e},l-1}_1>...>u^{\mathrm{e},l-1}_{k-l}$, hence a contradiction. Therefore, $u^{\mathrm{e},l-1}_1\geq 0$, and the recursive expressions of $(u^{\mathrm{e},l-1}_i)_i$ now imply that $0\geq u^{\mathrm{e},l-1}_1\geq u^{\mathrm{e},l-1}_2\geq ...\geq u^{\mathrm{e},l-1}_{k-l}$.

rearranging, we get

$$\Big[1-\delta\frac{x}{N-1}-\delta(1-x)\Big]u_1^{\mathrm{e},l-1}+(1-\delta)\sum_{i=2}^{k-l-2}u_i^{\mathrm{e},l-1}+\Big[1-\delta\Big(1-(1-x)\frac{k-l-1}{N-1}\Big)\Big]u_{k-l-1}^{\mathrm{e},l-1}$$
$$=\delta x\frac{k+l-1}{N-1}u_{k-l}^{\mathrm{e},l-1}>\delta x\frac{k+l-1}{N-1}u_{k-l-1}^{\mathrm{e},l-1}$$

Therefore, as $u_1^{\mathrm{e},l-1}<u_{k-l-1}^{\mathrm{e},l-1}$, rearranging implies that

$$\Big[2-\delta\Big(1+\frac{k+l}{N-1}\Big)\Big]u_{k-l-1}^{\mathrm{e},l-1}+(1-\delta)\sum_{i=2}^{k-l-2}u_i^{\mathrm{e},l-1}>0,$$

which is a contradiction, as $u_1^{\mathrm{e},l-1}<...<u_{k-l}^{\mathrm{e},l-1}\leq 0$. Consequently, $u_{k-l}^{\mathrm{e},l-1}\leq u_{k-l-1}^{\mathrm{e},l-1}$.

- consider the second inequality and suppose by contradiction that $u_{k-l}^{\mathrm{e},l-1}>u_{k-l+1}^{\mathrm{e},l-1}$. Using the recursive expression of $u_{k-l+1}^{\mathrm{e},l-1}$, this implies that $u_{k-l+2}^{\mathrm{e},l-1}<u_{k-l+1}^{\mathrm{e},l-1}<0$, and by induction that $0>u_{k-1}^{\mathrm{e},l-1}$. However, we know from the above computations that $u_i^{\mathrm{e},l-1}>0$ for any $i\geq k-1$, and thus in particular, $u_{k-1}^{\mathrm{e},l-1}>0$, which contradicts the above implication. Hence, $u_{k-l}^{\mathrm{e},l-1}\leq u_{k-l+1}^{\mathrm{e},l-1}$.

Therefore,

$$s+b+\delta\Big(\frac{k-l-1}{N-1}u_{k-l-1}^{\mathrm{e},l-1}+\frac{k+l-1}{N-1}u_{k-l}^{\mathrm{e},l-1}\Big)\geq s+b+\delta\frac{N-2}{N-1}u_{k-l}^{\mathrm{e},l-1},$$

and using the recursive expression of $u_{k-l}^{\mathrm{e},l-1}$,[40]

$$\Big[1-\delta\big[1-x(1-\lambda)\big]\frac{N-2}{N-1}\Big]u_{k-l}^{\mathrm{e},l-1}\geq-(1-\lambda)x(s-b).$$

As a consequence,

$$s+b+\delta\Big(\frac{k-l-1}{N-1}u_{k-l-1}^{\mathrm{e},l-1}+\frac{k+l-1}{N-1}u_{k-l}^{\mathrm{e},l-1}\Big)\geq s+b-\frac{\delta x(1-\lambda)(N-2)}{N-1-\delta\big[1-x(1-\lambda)\big](N-2)}(s-b)$$
$$\geq s+b-\frac{k-1}{k+1}(s-b)>0$$

Hence inequality (26) holds, as was to be shown.

(b). We now show existence, i.e. that for s/b sufficiently close to 1 there can be no profitable deviation from full entrenchment. The argument is analogous to the one just used.

[40]Namely,

$$u_{k-l}^{\mathrm{e},l-1}=-(1-\lambda)x(s-b)+\delta(1-x)\Big[\frac{k-l-1}{N-1}u_{k-l-1}^{\mathrm{e},l-1}+\frac{k+l+1}{N-1}u_{k-l}^{\mathrm{e},l-1}\Big]+\delta x\lambda\Big[\frac{k-l}{N-1}u_{k-l}^{\mathrm{e},l-1}+\frac{k+l-2}{N-1}u_{k-l+1}^{\mathrm{e},l-1}\Big]$$

In order to alleviate the notation, we again omit the superscript and the dependence on η and simply write V for the value function and u for its first difference.

The deviation differential payoff from full-entrenchment to entrenchment at a lower level in $M = N-1$ whenever the minority candidate is more talented writes

$$\Delta \equiv (1-\lambda)\Big[s - b - \delta \frac{N-2}{N-1} u_{N-2}\Big]$$

Explicit computation with (2)-(3) yield:

$$u_{N-2} = \delta(1-x\lambda)\frac{N-2}{N-1}u_{N-2} + \delta x\lambda\Big[\frac{N-3}{N-1}u_{N-3} + \frac{1}{N-1}u_{N-2}\Big]$$

and more generally for any $M \geq k$,

$$u_M = \delta(1-x\lambda)\Big[\frac{M}{N-1}u_M + \Big(1-\frac{M+1}{N-1}\Big)u_{M+1}\Big] + \delta x\lambda\Big[\frac{M-1}{N-1}u_{M-1} + \Big(1-\frac{M}{N-1}\Big)u_M\Big]$$

while for any $i \leq k-2$,

$$u_i = \delta(1-x\lambda)\Big[\frac{i-1}{N-1}u_{i-1} + \Big(1-\frac{i}{N-1}\Big)u_i\Big] + \delta x\lambda\Big[\frac{i-1}{N-1}u_i + \Big(1-\frac{i+1}{N-1}\Big)u_{i+1}\Big]$$

with

$$\Big[1-\delta(1-x\lambda)\Big]u_{k-1} = (1-2x\lambda)b + \delta(1-x\lambda)\Big[\frac{k-1}{N-1}u_k + \frac{k-2}{N-1}u_{k-2}\Big]$$

Summing up over all indices yields

$$\Big[1-\delta\Big(1-x\lambda\frac{N-2}{N-1}\Big)\Big]u_{N-2} + \Big[1-\delta\Big(1-\frac{x\lambda}{N-1}\Big)\Big]u_1 + (1-\delta)\sum_{i=2}^{N-3}u_i = (1-2x\lambda)b > 0 \tag{27}$$

Fix $b > 0$ and let $s = b$. The usual argument implies that $u_{N-2} > 0$. Indeed, if not, then the above equations imply by induction that $u_k \leq u_{k+1} \leq ... \leq u_{N-2} \leq 0$ and thus $0 \geq u_1 \geq u_2 \geq ... \geq u_{k-1}$, which yields to a contradiction with (27). Therefore, $u_{N-2} > 0$, and by induction again $u_k > u_{k+1} > ... > u_{N-2} > 0$. Hence the differential deviation payoff when the majority has size $N-2$ writes for $s = b$ as

$$\Delta = -(1-\lambda)\delta\frac{N-2}{N-1}u_{N-2} < 0$$

The result for $s = b$ obtains by noting that since $u_k > u_{k+1} > ... > u_{N-2} > 0$, the one-shot deviation when majority has size $N-1$ is the most profitable one-shot deviation from the full-entrenchment strategy. The result then extends to s/b in a neighbourhood of 1 by continuity.

Proof of claim (ii). The principal cannot expand the existence region of meritocracy by its interventions as the prospect of its overruling a majority's decision only scales down (by a strictly positive factor) the one-shot deviation differential payoff from meritocracy to entrenchment. Hence, under our assumption that the meritocratic equilibrium is selected whenever it exists, the principal fails to expand the region where one should expect meritocracy.

Whenever informed, the principal has a profitable (one-shot) deviation from no-intervention. Hence, if the principal cannot commit, majority members anticipate the principal steps in whenever informed. Proposition 4 then implies that an entrenched organization at best remains (canonically) entrenched or meritocratic, and otherwise goes super-entrenched – and most notably fully-entrenched for s/b in a non-empty neighbourhood of 1.

Hence in particular, for s/b sufficiently close to 1, the organization is fully-entrenched. Since the principal is only informed with probability strictly below 1, it cannot compensate all the "un-meritocratic" decisions made by the organizations. Hence, at any majority size $M \geq k+1$, the principal would be better off in terms of flow payoffs, if it could commit not to intervene.

By contrast, whenever the majority is tight, entrenchment would have prevailed, and so the principal may find it optimal to intervene. Fix a probability $\lambda \in (0,1)$ of the principal being informed, and consider any s/b sufficiently close to 1 such that, given λ, the unique equilibrium is full entrenchment. Suppose the principal values only ergodic efficiency. Then it would be better off (in terms of flow payoff when the majority size is k) committing not to intervene if and only if

$$N(N-1)(1-\lambda)xs > N(N-1)\nu^{\mathrm{e}}_{k+1}\frac{k+1}{N}xs, \qquad \text{i.e.} \qquad \lambda < 1-\nu^{\mathrm{e}}_{k+1}\frac{k+1}{N}$$

Hence, in terms of aggregate ergodic per-period welfare W, fixing s/b and a lower bound $\underline{\lambda}$ on the probability of the principal being informed such that the organization is super-entrenched for any $\lambda \geq \underline{\lambda}$. Then (a) the principal's objective W increases with $\lambda \geq \underline{\lambda}$, and (b) for any λ sufficiently low ($\lambda \geq \underline{\lambda}$), the principal would be better off if it could commit not to intervene.

More generally, fixing s/b, the principal's (ergodic aggregate) payoff increases with its probability λ of being informed on any interval such that the Pareto-dominating equilibrium remains unchanged. As the latter changes (a higher λ generating a higher level of entrenchment), the principal's payoff drops to a strictly lower value. Figure 4 thus depicts the principal's ergodic aggregate payoff as a functionits probability λ of being informed for fixed $s/b < \rho^{\mathrm{e}}$, assuming equilibrium selection and that the principal values only ergodic aggregate quality.

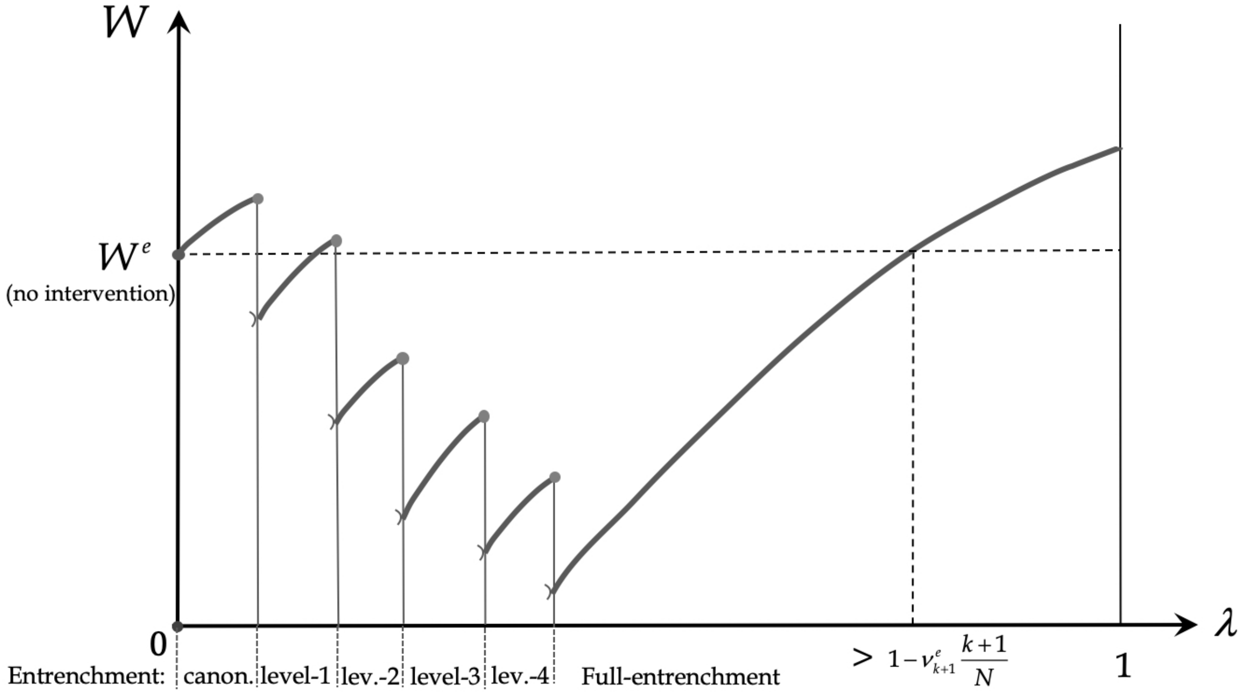

Figure 4: Principal's ergodic aggregate welfare as a function of λ (for $N = 12$).

I Proof of Proposition 5

Consider an entrenched organization. Because of the equilibrium selection (by Proposition 2, meritocracy thus prevails whenever it exists as an equilibrium), $s/b < \rho^{\mathrm{m}}$. Let $T \equiv \eta y$ denote equal the minimal expected bonus per member needed for the organization to move from entrenchment to meritocracy[41]. For the sake of exposition, we first assume that the principal does not value members' homophily benefits, and thus letting ξ be the cost of public funds[42], the principal's objective function writes as the ergodic welfare with per-period welfare given by $W = qS - \xi T$ [43]. Note that such an objective constitutes an *upper* bound on the admissible cost of a policy as (ergodic aggregate) homophily payoffs decrease when the organization goes from entrenchment to meritocracy (see Section 2.2.2). From previous computations on ergodic welfare, the (ergodic) efficiency gain from disentrenchment writes as $S^{\mathrm{m}} - S^{\mathrm{e}} = N(N-1)\nu^{\mathrm{e}}_{k+1}\dfrac{k+1}{N}x\dfrac{\tilde{s}}{1-\delta} > 0$. Rewarding quality is thus optimal for the principal if and

[41] Namely,

$$\frac{s^{+}(\eta, y)}{b} = \rho^{\mathrm{m}}, \qquad \text{i.e.} \qquad \eta y = \left(\frac{b}{s}\rho^{\mathrm{m}} - 1\right)\tilde{s} > 0$$

[42] The interpretation of ξ depends on the principal's welfare objective. If it is solely concerned with maximizing the (ergodic aggregate) quality of the organization, then ξ is the total cost of intervention, i.e. the sum of the payment and its shadow cost. By contrast, if the principal internalizes the "material" welfare of members, i.e. the sum of their quality payoffs and (possibly) rewards for quality (as opposed to their non-material welfare which consists of homophily benefits), then ξ is only the shadow cost of public funds.

[43] This objective may be interpreted as the limit of the main objective for $q, \xi \to \infty$.

only if

$$\xi\eta y N^2(\overline{x}+x) \leq N(N-1)\nu^{\mathrm{e}}_{k+1}\frac{k+1}{N}x\tilde{s}$$

where $N[\overline{x}+x]$ is the average number of talented members in a meritocratic organization, and ν^{e}_{k+1} the objective ergodic probability of majority size $k+1$ in the entrenched equilibrium (see Section 2.2.2). The above inequality rewrites as a condition on the administrative cost of public funds:[44]

$$\xi \leq \frac{(k+1)(N-1)}{N^2}.\frac{x\nu^{\mathrm{e}}_{k+1}}{\overline{x}+x}.\frac{\tilde{s}}{\eta y} = \frac{(k+1)(N-1)}{N^2}.\frac{x\nu^{\mathrm{e}}_{k+1}}{\overline{x}+x}.\frac{\frac{s}{b}}{\rho^{\mathrm{m}}-\frac{s}{b}}$$

Note that the RHS strictly increases with s/b and goes to $+\infty$ as s/b goes to ρ^{m} [45]. The result follows. The same argument applies if the principal's objective writes as $W = qS + B - \xi T$, yielding a higher threshold ρ_ξ (as $B^{\mathrm{m}} < B^{\mathrm{e}}$).

J Proof of Proposition 6

J.1 Proof of Proposition 6-(i): Affirmative action expands the existence region of meritocracy

Whenever a representation threshold is implemented, we refer to the "existence region of meritocracy" as the set of values of s/b for which there exists an equilibrium in which recruitments are meritocratic (i.e. a talented candidate is always recruited against a strictly less talented candidate). The result is (almost) immediate[46] for a representation threshold of 1. We thus focus on $R \geq 2$. [We provide in the Supplementary Materials explicit computations for the case $R = k-1$, which may be helpful to get the spirit of the proof in the general case.]

Consider a representation threshold of $R = k - l$ with $l \in \{1, ..., k-2\}$ and denote by

[44] By Inequality (16), a lower bound on the RHS of the above equation is given by

$$\frac{(k+1)(N-1)}{N^2}.\frac{x\nu^{\mathrm{e}}_{k+1}}{\overline{x}+x}.\frac{\tilde{s}}{\eta y} \geq \frac{(k+1)(N-1)^2}{(k-1)N^2}.\frac{x(1-2x)\nu^{\mathrm{e}}_{k+1}}{\overline{x}+x}.\frac{(1-\delta)}{\delta}$$

[45] The monotonicity of the RHS with respect to N is non-trivial. Namely, although the first two terms decrease with $N \geq 4$, so that $(k+1)(N-1)\nu^{\mathrm{e}}_{k+1}/N^2$ decreases with N, the comparative statics of ρ^{m} with respect to N are non-trivial. Nonetheless, for N large, the first two terms $(k+1)(N-1)\nu^{\mathrm{e}}_{k+1}/N^2$ are in $O(1/N)$, while for $\delta_0 < 1$, ρ^{m} is in $0(1)$. Therefore, the RHS is in $0(1/N)$ for N large, which is intuitive: the upper bound on the admissible cost of public funds is inversely proportional to the size of the organization, i.e. to the number of individuals to whom the bonus must be distributed.

[46] The argument is significantly shorter in this case than with $R \geq 2$ since the minority's value function in the canonical entrenched equilibrium writes as in the baseline model with no affirmative action (due to the conditioning on still being a member next period).

$\tilde{V}$ the value function under affirmative action with representation threshold R (omitting the superscript m). We will first show that the sequence $(\tilde{u}_i)_{i \geq k-1}$ satisfies at least one of the following assertions: (A_1) it decreases with i, or (A_2) it is always strictly negative; and that in particular $\tilde{u}_{k+l-1} < 0$ [47]. As in the baseline case, the monotonicity property (A_1) implies that the most tempting deviation from meritocracy to entrenchment is when the majority has size k and faces an untalented ingroup candidate and a talented outgroup candidate, while if (A_2) holds, then all deviations to entrenchment are non-profitable as they yield a deviation payoff equal to

$$-(s-b) + \delta\left[\left(1 - \frac{i}{N-1}\right)\tilde{u}_i + \frac{i-1}{N-1}\tilde{u}_{i-1}\right] < 0$$

By contrast, the sign of $\tilde{u}_{k+l-1}$ suggests there may be profitable deviations from meritocracy with ties broken in favor of the majority candidate to meritocracy with ties broken in favor of the minority candidate when s/b is high enough. (Lastly, because of discounting, there can be no profitable deviation consisting in voting an untalented minority candidate instead of a talented majority one.)

We first suppose by contradiction that $\tilde{u}_{k+l-1} \geq 0$. The usual induction argument relying on (4) then yields that $\tilde{u}_{k-1} > \tilde{u}_k > ... > \tilde{u}_{k+l-1} \geq 0$. Yet, summing as in the proof of Lemma 1, the above recursive expression for $\tilde{u}_{k+l-1}$ with (7) and (4) over indices k to $k+l-2$, and rearranging, yields on the LHS a weighted sum of $\tilde{u}_{k-1}, ..., \tilde{u}_{k+l-1}$ which is strictly positive, while on the RHS:

$$-xs - (1-x)b + (1-2x)b + \delta(1-x)\frac{k-2}{N-1}\tilde{u}_{k-2} = -x(s+b) + \delta(1-x)\frac{k-2}{N-1}\tilde{u}_{k-2},$$

and so $\tilde{u}_{k-2} > 0$. Summing (6) in $k-2$ to the above sum, and rearranging, yields on the LHS a weighted sum of $\tilde{u}_{k-1}, ..., \tilde{u}_{k+l-1}$ which is strictly positive, and on the RHS:

$$-x(s+b) + \delta(1-x)\frac{k-3}{N-1}\tilde{u}_{k-3},$$

Hence, $\tilde{u}_{k-3} > 0$, and by repeating this argument, $\tilde{u}_i > 0$ for any $i \in \{k-l-1, ..., k+l-1\}$. Yet summing the above recursive expressions of $\tilde{u}_{k-l-1}$ and $\tilde{u}_{k+l-1}$ together with (4)-(6)-(7) for $i \in \{k-l, ..., k+l-2\}$, yields after rearranging, on the LHS a weighted sum of all $\tilde{u}_i$ which is strictly positive, while on the RHS: $-x(s+b) + xs - (1-x)b = -b < 0$, which is a contradiction. Consequently, $\tilde{u}_{k+l-1} < 0$.

[47] A sketch of the proof is as follows: suppose by contradiction that $\tilde{u}_{k+l-1} \geq 0$, then by induction using (4), $\tilde{u}_{k-1} > ... > \tilde{u}_{k+l-1} \geq 0$. Yet summing the recursive expession of $\tilde{u}_i$ for $i \in \{k-1, ..., k+l-1\}$ implies that $\tilde{u}_{k-2} \geq 0$ as the sum of flow differential payoffs is equal to $-xs - (1-x)b + (1-2x)b = -x(s+b) < 0$. Using repeatedly the same argument gives that $u_i \geq 0$ for all $i \in \{k-l-1, ..., k+l-1\}$. A contradiction then obtains by summing the recursive expressions of u_i over the indices $i \in \{k-l-1, ..., k+l-1\}$, and noting that the sum of flow differential payoffs is equal to $-b < 0$.

In order to show that the sequence $(\tilde{u}_i)_{i\geq k-1}$ satisfies either (A_1) or (A_2) (or both), we proceed by induction considering the lowest index i^- such that $\tilde{u}_i < 0$ for any $i \geq i^-$. We first note that if $i^- \geq k$, then (4) brings by induction that[48]

$$\tilde{u}_{k+l-1} < \tilde{u}_{k+l-2} < ... < \tilde{u}_{i^-+1} < \tilde{u}_{i^-} < 0 < \tilde{u}_{i^--1} < \tilde{u}_{i^--2} < ... < \tilde{u}_{k-1},$$

which yields that (A_1) holds. If $i^- \leq k-1$, then (A_2) holds. By contrast, in the baseline setting without affirmative action, the sequence $(u_i)_{i\geq k-1}$ is positive for any i and decreases with i.

Consequently, in order to show that the existence region of meritocracy expands for low s/b with affirmative action, we only need to consider deviations from meritocracy to entrenchment when the majority is tight and faces an untalented ingroup candidate and a talented outgroup one, and show that the condition for non-profitability is looser with affirmative action than in the baseline setting.

Explicit computations yield[49]

$$\begin{cases} \tilde{u}_{k+l-1} = -xs - (1-x)b + \delta x\left[\dfrac{k-l}{N-1}\tilde{u}_{k+l-1} + \dfrac{k+l-2}{N-1}\tilde{u}_{k+l-2}\right] \\ \tilde{u}_{k-l-1} = xs - (1-x)b + \delta x\left[\dfrac{k-l-1}{N-1}\tilde{u}_{k-l-1} + \dfrac{k+l-1}{N-1}\tilde{u}_{k-l}\right] \end{cases} \tag{28}$$

Thus using (4) in $k+l-1$ and (6) in $k-l-1$, together with the fact that $u_i \geq 0$ for all i in

[48] The inequalities $\tilde{u}_{k+l-1} < \tilde{u}_{k+l-2} < ... < \tilde{u}_{i^-+1} < \tilde{u}_{i^-} < 0$ can be established by induction using the recursive expressions of the $\tilde{u}_i$ from $i = i^-$ up to $i = k+l-2$.

[49] By definition of affirmative action with representation threshold R, in any equilibrium

$$\begin{cases} \tilde{V}_{k+l} = \overline{x}s + \delta\left[\dfrac{k+l-1}{N-1}\tilde{V}_{k+l-1} + \dfrac{k-l}{N-1}\tilde{V}_{k+l}\right] \\ \tilde{V}_{k-l-1} = \overline{x}s + \delta\left[\dfrac{k-l-1}{N-1}\tilde{V}_{k-l-1} + \dfrac{k+l}{N-1}\tilde{V}_{k-l}\right] \end{cases}$$

Hence, in the meritocratic equilibrium,

$$\begin{cases} \tilde{u}_{k+l-1} = -xs - (1-x)b + \delta x\left[\dfrac{k-l}{N-1}\tilde{u}_{k+l-1} + \dfrac{k+l-2}{N-1}\tilde{u}_{k+l-2}\right] \\ \tilde{u}_{k-l-1} = xs - (1-x)b + \delta x\left[\dfrac{k-l-1}{N-1}\tilde{u}_{k-l-1} + \dfrac{k+l-1}{N-1}\tilde{u}_{k-l}\right] \end{cases}$$

the baseline setting, one gets[50]

$$\begin{cases} \left[1-\delta x\dfrac{k-l}{N-1}\right](\tilde{u}_{k+l-1}-u_{k+l-1}) < -xs-(1-x)b+\delta x\dfrac{k+l-2}{N-1}(\tilde{u}_{k+l-2}-u_{k+l-2}) \\ \left[1-\delta x\dfrac{k-l-1}{N-1}\right](\tilde{u}_{k-l-1}-u_{k-l-1}) < xs-(1-x)b+\delta x\dfrac{k+l-2}{N-1}(\tilde{u}_{k-l}-u_{k-l}) \end{cases}$$

Therefore, using (4) in $k+l-2$ and (6) in $k-l$, one gets

$$\begin{cases} \left[1-\delta x\dfrac{k-l+1}{N-1}-\delta(1-x)\dfrac{k+l-2}{N-1}-\delta(1-x)\dfrac{k-l}{N-1}\dfrac{\delta x\frac{k+l-2}{N-1}}{1-\delta x\frac{k-l}{N-1}}\right](\tilde{u}_{k+l-2}-u_{k+l-2}) \\ \quad < \dfrac{\delta(1-x)\frac{k-l}{N-1}}{1-\delta x\frac{k-l}{N-1}}\big[-xs-(1-x)b\big]+\delta x\dfrac{k+l-3}{N-1}(\tilde{u}_{k+l-3}-u_{k+l-3}) \\ \left[1-\delta x\dfrac{k-l}{N-1}-\delta(1-x)\dfrac{k+l-1}{N-1}-\delta(1-x)\dfrac{k-l-1}{N-1}\dfrac{\delta x\frac{k+l-1}{N-1}}{1-\delta x\frac{k-l-1}{N-1}}\right](\tilde{u}_{k-l}-u_{k-l}) \\ \quad < \dfrac{\delta(1-x)\frac{k-l-1}{N-1}}{1-\delta x\frac{k-l-1}{N-1}}\big[xs-(1-x)b\big]+\delta x\dfrac{k+l-2}{N-1}(\tilde{u}_{k-l+1}-u_{k-l+1}) \end{cases}$$

We begin by noting that

$$\frac{k-l}{N-1}\left[1-\delta x\frac{k-l-1}{N-1}\right] > \frac{k-l-1}{N-1}\left[1-\delta x\frac{k-l}{N-1}\right],$$

[50] Note that the omitted terms write for the first equation as

$$-\delta(1-x)\left[\frac{k+l-1}{N-1}u_{k+l-1}+\frac{k-l-1}{N-1}u_{k+l}\right],$$

which is thus proportional to $(-b)$ (see proof of Lemma 1 for details). Similarly for the second equation.

and[51]

$$\delta x\delta(1-x)\left(\frac{k-l}{N-1}\right)^2\frac{k+l-2}{N-1}\left[1-\delta x\frac{k-l-1}{N-1}\right]$$
$$> \delta x\delta(1-x)\frac{k-l+1}{N-1}\frac{k-l-1}{N-1}\frac{k+l-1}{N-1}\left[1-\delta x\frac{k-l}{N-1}\right]-\frac{\delta x}{N-1},$$

Hence, we have that[52]

$$\frac{k-l+1}{N-1}\left[1-\delta x\frac{k-l}{N-1}-\delta(1-x)\frac{k+l-1}{N-1}-\delta(1-x)\frac{k-l-1}{N-1}\frac{\delta x\frac{k+l-1}{N-1}}{1-\delta x\frac{k-l-1}{N-1}}\right]$$
$$> \frac{k-l}{N-1}\left[1-\delta x\frac{k-l+1}{N-1}-\delta(1-x)\frac{k+l-2}{N-1}-\delta(1-x)\frac{k-l}{N-1}\frac{\delta x\frac{k+l-2}{N-1}}{1-\delta x\frac{k-l}{N-1}}\right]$$

By downward (resp. upward) induction on $(\tilde{u}_i-u_i)$ for $i\geq k$ (resp. for $i\leq k-2$), we will get that

$$C_1(\tilde{u}_{k-1}-u_{k-1})<-C_2xs-C_3(1-x)b<0 \tag{29}$$

where C_1, C_2 and C_3 are strictly positive constants that depend on the parameters k, l and x. We detail the induction argument. Using (4)-(6), we obtain two sequences $(a_j)_{0\leq j\leq l-2}$ and

[51]To see this, we observe that: $(k-l)(k+l-2)=(k-l+1)(k+l-1)-(2k-1)$, and as a consequence, using the above inequality,

$$\left(\frac{k-l}{N-1}\right)^2\frac{k+l-2}{N-1}\left[1-\delta x\frac{k-l-1}{N-1}\right]$$
$$> \frac{k-l+1}{N-1}\frac{k-l-1}{N-1}\frac{k+l-1}{N-1}\left[1-\delta x\frac{k-l}{N-1}\right]-\frac{k-l}{N-1}\left[1-\delta x\frac{k-l-1}{N-1}\right]\frac{1}{N-1},$$

The inequality thus obtains using that $\delta(1-x)\frac{k-l}{N-1}<1-\delta x\frac{k-l}{N-1}$.

[52]Note that

$$\frac{k-l+1}{N-1}\left[1-\delta(1-x)\frac{k+l-1}{N-1}\right]=\frac{k-l}{N-1}\left[1-\delta(1-x)\frac{k+l-2}{N-1}\right]+\frac{1-\delta(1-x)}{N-1}$$

$(b_j)_{0\leq j\leq l-2}$ such that for any $j \leq l-2$,

$$\begin{cases} a_j\left(\tilde{u}_{k+j}-u_{k+j}\right) \\ \quad < -[xs+(1-x)b]\dfrac{\delta(1-x)\dfrac{k-l}{N-1}}{1-\delta x\dfrac{k-l}{N-1}}\displaystyle\prod_{n=j+1}^{l-2}\left(\frac{\delta(1-x)}{a_n}\frac{k-n-1}{N-1}\right)+\delta x\frac{k+j-1}{N-1}\left(\tilde{u}_{k+j-1}-u_{k+j-1}\right) \\ b_j\left(\tilde{u}_{k-j-2}-u_{k-j-2}\right) \\ \quad < [xs-(1-x)b]\dfrac{\delta(1-x)\dfrac{k-l-1}{N-1}}{1-\delta x\dfrac{k-l-1}{N-1}}\displaystyle\prod_{n=j+1}^{l-2}\left(\frac{\delta(1-x)}{b_n}\frac{k-n-2}{N-1}\right)+\delta x\frac{k+j}{N-1}\left(\tilde{u}_{k-j-1}-u_{k-j-1}\right) \end{cases}$$

where

$$\begin{cases} a_{j-1}=1-\delta x\dfrac{k-j}{N-1}-\delta(1-x)\dfrac{k+j-1}{N-1}-\delta(1-x)\dfrac{k-j-1}{N-1}\dfrac{\delta x\dfrac{k+j-1}{N-1}}{a_j} \\ b_{j-1}=1-\delta x\dfrac{k-j-1}{N-1}-\delta(1-x)\dfrac{k+j}{N-1}-\delta(1-x)\dfrac{k-j-2}{N-1}\dfrac{\delta x\dfrac{k+j}{N-1}}{b_j} \end{cases}$$

We first note that by induction[53]

$$\forall j\leq l-1,\qquad \frac{\delta(1-x)}{a_j}\frac{k-j-1}{N-1}<1,\qquad \text{and}\qquad \frac{\delta(1-x)}{b_j}\frac{k-j-2}{N-1}<1 \tag{30}$$

Hence, using (7), we have that the coefficient C_1 in (29) is given by

$$1-\delta(1-x)-\frac{\delta(1-x)}{a_0}\frac{k-1}{N-1}\delta x\frac{k-1}{N-1}-\frac{\delta(1-x)}{b_0}\frac{k-2}{N-1}\delta x\frac{k}{N-1}>1-\delta>0$$

Using (7) further implies that the coefficient C_3 in (29) is strictly positive. We then show by downward induction on j that for any $j\leq l-1$,

$$\frac{1}{a_j}\frac{k-j-1}{N-1}>\frac{1}{b_j}\frac{k-j-2}{N-1},$$

which will yield that $C_2>0$. The initialization ($j=l-1$) derives from the observation in footnote 53 (the case $j=l-2$ has also been established above). As for the induction, i.e. to

[53] The initialization with $j=l-1$ stems from the observation that

$$\delta(1-x)\frac{k-l}{N-1}<1-\delta x\frac{k-l}{N-1},\qquad \text{and}\qquad \delta(1-x)\frac{k-l-1}{N-1}<1-\delta x\frac{k-l-1}{N-1}$$

Moreover,

$$\delta(1-x)\frac{k-l}{N-1}\left[1-\delta x\frac{k-l-1}{N-1}\right]>\delta(1-x)\frac{k-l-1}{N-1}\left[1-\delta x\frac{k-l}{N-1}\right],\qquad \text{i.e.}\ \frac{1}{a_{l-1}}\frac{k-l}{N-1}>\frac{1}{b_{l-1}}\frac{k-l-1}{N-1}$$

show that $a_{j-1}(k-j-1) < b_{j-1}(k-j)$, we note that for any $j \geq 0$, the induction hypothesis implies that[54]

$$\frac{k-j-1}{N-1}\frac{k+j-1}{N-1}\frac{1}{a_j}\frac{k-j-1}{N-1} > \frac{k-j}{N-1}\frac{k+j}{N-1}\frac{1}{b_j}\frac{k-j-2}{N-1} - \frac{1}{a_j}\frac{k-j-1}{N-1}\frac{1}{N-1}$$

and thus, using (30),

$$\delta x\delta(1-x)\frac{k-j-1}{N-1}\frac{k+j-1}{N-1}\frac{1}{a_j}\frac{k-j-1}{N-1} > \delta x\delta(1-x)\frac{k-j}{N-1}\frac{k+j}{N-1}\frac{1}{b_j}\frac{k-j-2}{N-1} - \frac{\delta x}{N-1}$$

Therefore, using the recursive expression of a_{j-1} and b_{j-1}, we have that

$$a_{j-1}(k-j-1) < b_{j-1}(k-j) - \frac{1-\delta}{N-1} < b_{j-1}(k-j),$$

as was to be shown.

This in turn implies that $(\tilde{u}_k - u_k) < 0$. Therefore, the non-profitability conditions for deviations from meritocracy to entrenchment is (strictly) looser with a representation threshold R than without. Moreover, since C_2 and C_3 are strictly positive, and since the omitted negative terms are all proportional to $(-b)$ (see footnote 50), the existence region of meritocracy expands downward (i.e. for low values of s/b).[55]

For s/b sufficiently high, meritocracy with reverse favoritism is an equilibrium: the majority always picks the most talented candidate and breaks ties in favor of the minority candidate. Let $b = 0 < s$. We first note that in the unconstrained meritocratic equilibrium, this implies that $u_i = 0$ for any $i \in \{1, ..., N-2\}$. The above computations then apply, switching the weights $1-x$ and x (except for the flow payoffs of $\tilde{u}_{k+l-1}$ and $\tilde{u}_{k-l-1}$ which remain respectively given by $-xs$ and xs). Hence $\tilde{u}_i < 0$ for any $i \geq k-1$. Consequently, the deviation differential payoff from reverse-favoritism meritocracy to standard-favoritism meritocracy at majority size M is given by

$$\delta\Big(\frac{M-1}{N-1}\tilde{u}^{\mathrm{m}}_{M-1} + \frac{N-1-M}{N-1}\tilde{u}^{\mathrm{m}}_{M}\Big) < 0,$$

which yields the result. By contrast, the same argument implies that meritocracy with stan-

[54]Indeed, we have that $(k-j-1)(k+j-1) = (k-j)(k+j) - (2k-1)$, and

$$\frac{k-j}{N-1}\left[1-\delta(1-x)\frac{k+j}{N-1}\right] = \frac{k-j-1}{N-1}\left[1-\delta(1-x)\frac{k+j-1}{N-1}\right] + \frac{1-\delta(1-x)}{N-1}$$

[55]Moreover, since either (A_1) or (A_2) hold, we have by monotonicity with respect to b and s that for any value of the ratio s/b such that meritocracy exists, then for any higher value of the ratio, there can be no profitable deviations towards entrenchment (i.e. un-meritocratic decisions are always unprofitable). This establishes that meritocratic decisions – with ties broken either in favor of the majority or the minority – happen on a half-line, i.e. for any s/b sufficiently high.

dard favoritism is no longer an equilibrium for s/b sufficiently high.[56]

Non-ergodic welfare comparison. The same computations as in the proof of Proposition 2 (see Online Appendix E) apply. Therefore, whenever meritocracy and entrenchment coexist, at any majority size the meritocratic equilibrium is preferred to the entrenched equilibrium by (current) majority members. Building on analogous computations, it can be shown that the same preference also holds in several cases for all (current) minority members. By mimicking the argument in Online Appendix E – as well as in Online Appendix H –, we have that $\tilde{V}_i^{\mathrm{m}} \geq \tilde{V}_i^{\mathrm{e}}$ for any $i \leq k-1$ if

$$s + b + \delta\Big(\frac{k}{N-1}\tilde{u}_{k-1}^{\mathrm{m}} + \frac{k-2}{N-1}\tilde{u}_{k-2}^{\mathrm{m}}\Big) > 0, \tag{31}$$

Hence this inequality clearly holds whenever δ is small. In addition, we show in the Supplementary Materials that this inequality holds in several other cases. This dominance in terms of non-ergodic welfare motivates the welfare analysis of Proposition 6-(ii).

J.2 Proof of Proposition 6-(ii)

Let $N \geq 4$ and $1 \leq l \leq k-1$. The ergodic aggregate efficiency of a canonically entrenched organization under laissez-faire and a meritocratic one under affirmative action with representation threshold l write respectively:

$$\begin{cases} S^{\mathrm{e}} = N(N-1)\Big[\dfrac{k+1}{N}\nu_{k+1}^{\mathrm{e}}\overline{x} + \Big(1 - \dfrac{k+1}{N}\nu_{k+1}^{\mathrm{e}}\Big)(\overline{x}+x)\Big]\tilde{s} \\[2ex] S^{\mathrm{m,AA}} = N(N-1)\Big[\dfrac{l}{N}\nu_{N-l}^{\mathrm{m,AA}}\overline{x} + \Big(1 - \dfrac{l}{N}\nu_{N-l}^{\mathrm{m,AA}}\Big)(\overline{x}+x)\Big]\tilde{s} \end{cases}$$

and thus:

$$S^{\mathrm{m,AA}} - S^{\mathrm{e}} = N(N-1)\Big[\frac{k+1}{N}\nu_{k+1}^{\mathrm{e}} - \frac{l}{N}\nu_{N-l}^{\mathrm{m,AA}}\Big]x\tilde{s}$$

Explicit computations (see Lemma 2 and its proof in Section F) yield:

$$\begin{cases} \nu_{k+1}^{\mathrm{e}}\Big[1 + \displaystyle\sum_{i=1}^{k-1}\Big(\frac{1-x}{x}\Big)^i\prod_{j=1}^{i}\frac{k-j}{k+1+j}\Big] = 1 \\ \nu_{N-l}^{\mathrm{m,AA}}\Big[1 + \displaystyle\sum_{i=1}^{k-l-1}\Big(\frac{x}{1-x}\Big)^i\prod_{j=1}^{i}\frac{N-l+1-j}{l+j} + \Big(\frac{x}{1-x}\Big)^{k-l}\frac{k+1}{N}\prod_{j=1}^{k-l-1}\frac{N-l+1-j}{l+j}\Big] = 1 \end{cases}$$

[56] Considering $b = 0 < s$, and observing that in the unconstrained meritocratic equilibrium, $u_i = 0$ for any $i \in \{1, ..., N-2\}$ and using the above computations in order to get that $\tilde{u}_i < 0$.

Consequently, $S^{\mathrm{m,AA}} - S^{\mathrm{e}}$ has same sign as

$$(k+1)\left[1+\sum_{i=1}^{k-l-1}\left(\frac{x}{1-x}\right)^i\prod_{j=1}^{i}\frac{N-l+1-j}{l+j}+\left(\frac{x}{1-x}\right)^{k-l}\frac{k+1}{N}\prod_{j=1}^{k-l-1}\frac{N-l+1-j}{l+j}\right]$$
$$-\,l\left[1+\sum_{i=1}^{k-1}\left(\frac{1-x}{x}\right)^i\prod_{j=1}^{i}\frac{k-j}{k+1+j}\right]$$

We then note that the above expression is strictly negative for x in a neighbourhood of 0, and strictly positive for x in a neighbourhood of 1. Moreover, since $x/(1-x)$ (resp. $(1-x)/x$) strictly increases (resp. decreases) with $x \in (0, 1/2)$, there exists a unique $x_{\mathrm{AA}}(l) \in (0, 1/2]$ such that for any $x < x_{\mathrm{AA}}(l)$ (resp. $x > x_{\mathrm{AA}}(l)$), the above expression is strictly negative (resp. positive).

Lastly, we note that by construction, $x_{\mathrm{AA}}(l)$ is such that

$$(k+1)\left[1+\sum_{i=1}^{k-l-1}\left(\frac{x_{\mathrm{AA}}(l)}{1-x_{\mathrm{AA}}(l)}\right)^i\prod_{j=1}^{i}\frac{N-l+1-j}{l+j}+\left(\frac{x_{\mathrm{AA}}(l)}{1-x_{\mathrm{AA}}(l)}\right)^{k-l}\frac{k+1}{N}\prod_{j=1}^{k-l-1}\frac{N-l+1-j}{l+j}\right]$$
$$=l\left[1+\sum_{i=1}^{k-1}\left(\frac{1-x_{\mathrm{AA}}(l)}{x_{\mathrm{AA}}(l)}\right)^i\prod_{j=1}^{i}\frac{k-j}{k+1+j}\right]$$

The LHS in the above equation strictly decreases with l for any given x fixed, and strictly increases with x for any fixed l. By contrast, the RHS strictly increases with l for any fixed x, and strictly decreases with x for any fixed l. Hence $x_{\mathrm{AA}}(l)$ strictly increases with l.

K Proof of Proposition 7

We use a fixed-point argument to prove the existence of a class of equilibria characterized by a weakly decreasing decision rule $(\Delta_M)_M$ [57]. Let $\overline{u}$ be given by

$$\overline{u} \equiv \frac{1}{1-\delta}\Big(\mathbb{E}\big[(s+b)\mathbf{1}\{\hat{s}-s \leq b\}\big]+\mathbb{E}\big[\hat{s}\mathbf{1}\{\hat{s}-s > b\}\big]\Big)$$

Note that $(\mathbb{E}\big[(s+b)\mathbf{1}\{\hat{s}-s \leq b\}\big]+\mathbb{E}\big[\hat{s}\mathbf{1}\{\hat{s}-s > b\}\big])$ is the highest flow payoff a majority member can guarantee, and consequently, $\overline{u}$ represents an upper bound on the majority's expected utility from a recruitment (i.e. its expected utility in the absence of control consideration). We define K as the set of sequences $(u_M)_{M\in\{k-1,...,N-2\}}$ such that (i) for any M, $u_M \in [0, \overline{u}]$ and (ii) the sequence $(u_M)_M$ is weakly decreasing. By construction, the set K is non-empty, compact and convex.

As earlier, let $\{V_i\}$ denote the value functions and $V \equiv (V_1, ..., V_{N-1})$. For $i \in \{k-1, ..., N-2\}$, let $u_i \equiv V_{i+1} - V_i$. In the equilibria we look for, whenever the majority has

[57] We thus focus on equilibria such that the decision rule only depends on the majority size.

size $M \in \{k, ..., N-1\}$, it favors a majority candidate with (discounted) talent s against a minority candidate with (discounted) talent $\hat{s}$ if and only if[58]

$$\hat{s} + \delta\Big[\frac{M-1}{N-1}V_{M-1} + \Big(1 - \frac{M-1}{N-1}\Big)V_M\Big] \leq s + b + \delta\Big[\frac{M}{N-1}V_M + \Big(1 - \frac{M}{N-1}\Big)V_{M+1}\Big],$$

i.e. if and only if

$$\hat{s} - s \leq b + \delta\Big[\frac{M-1}{N-1}u_{M-1} + \Big(1 - \frac{M}{N-1}\Big)u_M\Big]$$

We denote by $\overline{s} \in [b, +\infty)$ the lowest real number such that $\mathbb{P}(\hat{s} - s \leq \overline{s}) = 1$ if it exists, and let $\overline{s} = +\infty$ otherwise. We first consider the "decision-rule" (cutoff) mapping $D : K \longrightarrow \big[0, \min(b + \delta\overline{u}, \overline{s})\big]^k, u \longmapsto (D_M)_{M\in\{k,\ldots,N-1\}}$, where

$$D_M(u) \equiv \begin{cases} b + \delta\Big[\frac{M-1}{N-1}u_{M-1} + \Big(1 - \frac{M}{N-1}\Big)u_M\Big] & \text{if } b + \delta\Big[\frac{M-1}{N-1}u_{M-1} + \Big(1 - \frac{M}{N-1}\Big)u_M\Big] < \overline{s} \\ \overline{s} & \text{otherwise} \end{cases}$$

Taking $V_{k-1} \geq 0$ as fixed, we consider the "value-function" mapping T defined as $T : \big[0, +\infty\big]^k \times \big[b, \overline{s}\big]^k \longrightarrow \big[0, +\infty\big]^k, ((V_M)_M, (\Delta_M)_M) \longmapsto (T_M)_M$, where

$$\begin{aligned} T_M(V, \Delta) \equiv\ & \mathbb{E}\big[(s+b)\mathbf{1}\{\hat{s} - s \leq \Delta_M\}\big] + \delta\mathbb{P}(\hat{s} - s \leq \Delta_M)\Big[\frac{M}{N-1}V_M + \Big(1 - \frac{M}{N-1}\Big)V_{M+1}\Big] \\ & + \mathbb{E}\big[\hat{s}\mathbf{1}\{\hat{s} - s > \Delta_M\}\big] + \delta\mathbb{P}(\hat{s} - s > \Delta_M)\Big[\frac{M-1}{N-1}V_{M-1} + \Big(1 - \frac{M-1}{N-1}\Big)V_M\Big] \end{aligned}$$

In order to alleviate the notation, we define the functions h and h_1 as

$$\begin{cases} h(X) \equiv \mathbb{E}\big[(s+X)\mathbf{1}\{\hat{s} - s \leq X\}\big] + \mathbb{E}\big[\hat{s}\mathbf{1}\{\hat{s} - s > X\}\big] \\ h_1(X) \equiv X - h(X) \end{cases}$$

Fix $V_{k-1} \geq 0$. Given a sequence $u \equiv (u_M)_{M\in\{k-1,\ldots,N-2\}} \in K$, we define the sequence $V(u) \equiv (V_M)_{M\in\{k,\ldots,N-1\}}$ by upward induction by letting $V_M \equiv u_{M-1} + V_{M-1}$. Lastly, we define the mapping $\Upsilon : u \longmapsto \Upsilon(u)$ from K into itself by

$$\Upsilon_M(u) \equiv \min\Big\{T_{M+1}(V(u), D(u)) - T_M(V(u), D(u)), h(b)/(1-\delta)\Big\}$$

for any $M \in \{k-1, ..., N-2\}$ (with the convention that $T_{k-1}(V(u), D(u)) \equiv V_{k-1}$). While bounding above $\Upsilon(u)$ is necessary to the argument, it does not threaten the existence of an equilibrium: indeed, $h(b)$ is the highest flow payoff (quality and homophily) that a majority

[58]The assumption that ties are broken in favor of the majority candidate comes withouth loss of generality when vertical types are continuously distributed within each group.

member can guarantee[59]. Hence we have by construction that for any $u \in K$ and any $i \in \{k-1, ..., N-2\}$, $\Upsilon_i(u) \leq \overline{u}$. With an abuse of notation, we omit in the following the min operator.

We now check that the mapping Υ is well-defined, i.e. that $\Upsilon(u) \in K$ for any $u \in K$. Rearranging the above expression for $T_M(V(u), D(u))$ yields:

$$\begin{aligned}T_M(V(u), D(u)) = \mathbb{E}\big[s\mathbf{1}\{\hat{s} - s \leq D_M(u)\}\big] + \mathbb{E}\big[\hat{s}\mathbf{1}\{\hat{s} - s > D_M(u)\}\big] \\ + \mathbb{P}(\hat{s} - s \leq D_M(u))\Big[b + \delta\Big[\frac{M-1}{N-1}u_{M-1} + \Big(1 - \frac{M}{N-1}\Big)u_M\Big]\Big] \\ + \delta\Big[\frac{M-1}{N-1}V_{M-1} + \Big(1 - \frac{M-1}{N-1}\Big)V_M\Big]\end{aligned}$$

We thus distinguish two cases.

(A) If $D_M(u) < \overline{s}$ for all $M \geq k$, then[60]

$$\begin{aligned}T_M(V(u), D(u)) &= \mathbb{E}\big[s\mathbf{1}\{\hat{s} - s \leq D_M\}\big] + \mathbb{E}\big[\hat{s}\mathbf{1}\{\hat{s} - s > D_M\}\big] + \mathbb{P}(\hat{s} - s \leq D_M)D_M \\ &\quad + \delta\Big[\frac{M-1}{N-1}V_{M-1} + \Big(1 - \frac{M-1}{N-1}\Big)V_M\Big] \\ &= h(D_M) + \delta\Big[\frac{M-1}{N-1}V_{M-1} + \Big(1 - \frac{M-1}{N-1}\Big)V_M\Big] \qquad (32)\end{aligned}$$

Consequently, if $D_M(u) < \overline{s}$ [61], plugging the above expressions in the equality $\Upsilon_M(u) = T_{M+1}(V, D) - T_M(V, D)$, and using the expression of D_M as a function of u, yields

$$\begin{aligned}\Upsilon_M(u) &= h(D_{M+1}) - h(D_M) + \delta\Big[\frac{M-1}{N-1}u_{M-1} + \Big(1 - \frac{M}{N-1}\Big)u_M\Big] \\ &= h(D_{M+1}) + h_1(D_M) - b \qquad (33)\end{aligned}$$

Since $u \in K$, we have that (i) $u_M \geq 0$ for any M and thus by construction $D_M \geq b$, and (ii) the sequence $(u_M)_M$ is decreasing, and thus so is the sequence $(D_M)_M$. As a consequence, $D_M \geq D_{M+1} \geq b$.

Henceforth, we restrict our attention to joint distributions such that the functions h_1 and $(h - h_1)$ are strictly increasing over $[b, +\infty) \cap \mathrm{Supp}(\hat{s} - s)$ [62]. This set notably includes the set

[59] Indeed, for any joint distribution of types, the quantity

$$\mathbb{E}[(s+b)\mathbf{1}\{\hat{s} - s \leq X\}] + \mathbb{E}[\hat{s}\mathbf{1}\{\hat{s} - s > X\}]$$

decreases with $X \geq b$.

[60] Note that in this case the mapping T can be defined as $T : \big[0, V_{k-1} + k\overline{u}\big]^k \times \big[b, b + \overline{u}\big]^k \longrightarrow \big[0, V_{k-1} + k\overline{u}\big]^k$.

[61] By monotonicity (as $u \in K$), $D_M(u) < \overline{s}$ implies that $D_{M'} < \overline{s}$ for any $M' > M$.

[62] Note that $(h - h_1)$ being strictly increasing implies that h is strictly increasing, as $h(X) - h_1(X) = 2h(X) - X$.

of continuous joint symmetric distributions[63], as well as the case where the majority candidate has a fixed type $s \geq 0$ and the minority candidate a type $s + D$ where D is a (full support) random variable with a continuously differentiable distribution over $(-s, s)$ symmetric around 0.[64]

As a consequence, for any $u \in K$, $\Upsilon_M(u) \geq 0$ and the sequence $(\Upsilon_M(u))_{M \geq k}$ is decreasing as it inherits the monotonicity of the sequence $(D_M)_M$. Moreover, for any $M \geq k$,

$$\Upsilon_M(u) \leq h(D_M) + h_1(D_M) - b = \delta\Big[\frac{M-1}{N-1}u_{M-1} + \Big(1 - \frac{M}{N-1}\Big)u_M\Big] < \delta\frac{N-2}{N-1}u_{k-1} \leq \overline{u}$$

It thus remains to check that $\Upsilon_{k-1}(u) \geq \Upsilon_k(u)$. By monotonicity of h and $(h - h_1)$ and using the above computations, a sufficient condition for this inequality to hold writes as:

$$(1-\delta)V_{k-1} \leq h(b)$$

This condition imposes an upper bound on V_{k-1}. Recall that $h(b)$ is the highest flow payoff (quality and homophily) that a majority member can guarantee. Therefore, for any symmetric joint distribution of types, any (increasing and concave) equilibrium value function must satisfy $V_{k-1} < h(b)/(1-\delta)$. Hence assuming this inequality hold does not threaten the existence of an equilibrium. We thus fix in the following V_{k-1} such that the above inequality holds. Hence, under the above conditions, $\Upsilon(u) \in K$.

(B) We now consider the case where $\overline{s} < +\infty$ and $D_M(u) = \overline{s}$ for some M. (Note that as $u_M \leq \overline{u} < \infty$, the case $D_M(u) = \overline{s}$ can only arise when $\overline{s} < \infty$.)

We first note that, within the class of equilibria with $u \in K$ (and thus a decreasing sequence $(\Delta_M)_M$), $\Delta_k = \overline{s}$ implies that $\Delta_{k+1} < \overline{s}$. Hence, whenever the majority is not tight, it recruits

[63]Indeed, letting F be the marginal c.d.f. of s and $\hat{s}$, then

$$\forall \Delta > 0, \qquad h(\Delta) = \int_0^{\overline{s}} (s+\Delta)F(s+\Delta)dF(s) + \int_{\Delta}^{\overline{s}} \hat{s}F(\hat{s}-\Delta)dF(\hat{s}),$$

and thus, for any $\Delta \in (0, \overline{s})$,

$$h'(\Delta) = \int_0^{\overline{s}} F(s+\Delta)dF(s) + \int_0^{\overline{s}-\Delta}(s+\Delta)f(s+\Delta)dF(s) - \int_{\Delta}^{\overline{s}} \hat{s}f(\hat{s}-\Delta)dF(\hat{s}) = \int_0^{\overline{s}} F(s+\Delta)dF(s),$$

and thus $h'(\Delta) \in (1/2, 1)$ since $\displaystyle\int_0^{\overline{s}} F(s)dF(s) = 1/2$.

[64]Indeed, denoting by F the c.d.f. of D, we have for any $\Delta \in (0, \overline{s})$,

$$h(\Delta) = \int_{-s}^{\Delta}(s+\Delta)dF(D) + \int_{\Delta}^{s}(s+D)dF(D), \qquad \text{and thus} \qquad h'(\Delta) = F(\Delta) \in (1/2, 1)$$

a minority candidate with a strictly positive probability: $\Delta_M < \overline{s}$ for any $M \geq k+1$.[65]

Consequently, we only need to consider the case where $D_{k+1}(u) < D_k(u) = \overline{s} < \infty$ [66]. We first show that $\Upsilon_k(u) \in [\Upsilon_{k+1}(u), \overline{u}]$. By construction,

$$T_k(V(u), D(u)) = \mathbb{E}[s] + b + \delta\left[\frac{k}{N-1}V_k + \left(1 - \frac{k}{N-1}\right)V_{k+1}\right],$$

and thus, since $D_{k+1} < \overline{s}$ implies that $T_{k+1}(V, D)$ is given by (32),

$$\Upsilon_k(u) = h(D_{k+1}) - \mathbb{E}[s] - b$$

By monotonicity of the sequence $(D_M)_M$ and since the functions h and h_1 are increasing, we have that $\Upsilon_k(u) \geq \Upsilon_{k+1}(u)$. It thus remains to check that $\Upsilon_{k-1}(u) \geq \Upsilon_k(u)$. A sufficient condition for this inequality to hold writes as[67]

$$(1-\delta)V_{k-1} \leq \mathbb{E}[s] + b + \frac{k}{N-2}(\overline{s} - b)$$

[65] Indeed, suppose by contradiction that $\Delta_k = \Delta_{k+1} = \overline{s}$. Then, by construction,

$$u_k = \delta\left[\frac{k}{N-1}u_k + \frac{k-2}{N-1}u_{k+1}\right]$$

Since $u \in K$, this yields that $u_k = u_{k+1} = 0$, which contradicts the initial assumption as $b < \overline{s}$.

[66] Indeed, note that if $D_{k+1}(u) < \overline{s}$, then $D_{k+1}(\Upsilon(u)) < \overline{s}$ as

$$\begin{aligned} D_{k+1}(\Upsilon(u)) &< b + \delta\left[\frac{k}{N-1}\Big(h(D_{k+1}(u)) - \mathbb{E}[s] - b\Big) + \frac{k-2}{N-1}\Big(h(D_{k+2}(u) + h_1(D_{k+1}(u)) - b\Big)\right] \\ &< \left(1 - \delta\frac{N-2}{N-1}\right)b + \delta\left[\frac{k}{N-1}\big(h(D_{k+1}(u)) - \mathbb{E}[s]\big) + \frac{k-2}{N-1}D_{k+1}(u)\right] \\ &< \left(1 - \delta\frac{N-2}{N-1}\right)b + \delta\frac{N-2}{N-1}\overline{s} < \overline{s} \end{aligned}$$

[67] Indeed, a sufficient condition for $\Upsilon_{k-1}(u) \geq \Upsilon_k(u)$ is

$$2(\mathbb{E}[s] + b) - (1-\delta)V_{k-1} + \delta u_{k-1} \geq h\left(b + \delta\frac{N-2}{N-1}u_k\right) - \delta\frac{k-1}{N-1}u_k,$$

which by monotonicity of h and $h - h_1$ holds in particular if

$$\begin{aligned} & 2(\mathbb{E}[s] + b) - (1-\delta)V_{k-1} + \delta u_{k-1} \geq h\left(b + \delta\frac{N-2}{N-1}u_{k-1}\right) - \delta\frac{k-1}{N-1}u_{k-1}, \\ \text{i.e.} \quad & (1-\delta)V_{k-1} \leq 2(\mathbb{E}[s] + b) - h\left(b + \delta\frac{N-2}{N-1}u_{k-1}\right) + \delta\left(1 + \frac{k-1}{N-1}\right)u_{k-1} \end{aligned}$$

Hence, by monotonicity of $X \mapsto X - h(X)$ and since u_{k-1} must satisfy $\delta(N-2)/(N-1)u_{k-1} \geq (\overline{s} - b)$, a sufficient condition for this inequality to hold is

$$(1-\delta)V_{k-1} \leq 2(\mathbb{E}[s] + b) - h(\overline{s}) + (\overline{s} - b) + \frac{k}{N-2}(\overline{s} - b),$$

which yields the result as $h(\overline{s}) = \mathbb{E}[s] + \overline{s}$.

This second inequality is looser than the condition[68] in case (A) and is thus satisfied for $V_{k-1} \leq h(b)/(1-\delta)$ (which must be the case in any equilibrium as discussed above).

Therefore, fixing $V_{k-1} \in [0, h(b)/(1-\delta)]$, Υ is a well-defined continuous mapping from K into itself. By Brouwer's fixed point theorem, it admits a fixed point. This establishes existence.

We now show that any equilibrium characterized by a sequence of cut-offs $(\Delta_M)_{M \geq k}$ is such that (a) $\Delta_M > b$ for any $M \geq k$, and (b) the sequence $(\Delta_M)_M$ is strictly decreasing.

(a) We first argue that in any equilibrium, $\Delta_M > b$ for any $M \geq k$. We show this by downward induction. Suppose that $\Delta_{N-1} \leq b$. Then[69], this implies that $u_{N-2} \leq 0$, i.e. $V_{N-2} \geq V_{N-1}$. Hence the continuation payoff for a majority of size $N-1$ is bounded below by δV_{N-1}. By deviating from Δ_{N-1} to the value that maximizes the flow payoff, a majority with size $N-1$ gets a utility greater than

$$\max_X \left\{ \mathbb{E}[(s+b)\mathbf{1}\{\hat{s}-s \leq X\}] + \mathbb{E}[\hat{s}\mathbf{1}\{\hat{s}-s > X\}] \right\} + \delta V_{N-1}$$

Hence this would imply that

$$(1-\delta)V_{N-1} \geq \max_X \left\{ \mathbb{E}[(s+b)\mathbf{1}\{\hat{s}-s \leq X\}] + \mathbb{E}[\hat{s}\mathbf{1}\{\hat{s}-s > X\}] \right\}$$

which is a contradiction as the RHS is the highest attainable flow payoff (and $\delta > 0$). Therefore $V_{N-1} > V_{N-2}$, and thus $\Delta_{N-1} > b$. Suppose now that $V_{M'+1} > V_{M'}$ for any $M' \geq M$, and that $V_M \leq V_{M-1}$. Therefore, the continuation payoff for a majority of size M is bounded below by δV_M. Hence, by deviating from Δ_M to the value that maximizes the flow payoff, a majority with size M gets a utility greater than

$$\max_X \left\{ \mathbb{E}[(s+b)\mathbf{1}\{\hat{s}-s \leq X\}] + \mathbb{E}[\hat{s}\mathbf{1}\{\hat{s}-s > X\}] \right\} + \delta V_M,$$

which again leads to a contradiction. Consequently, $u_{M-1} > 0$, while $u_M > 0$ by the induction

[68] Indeed, for any joint distribution such that $(\hat{s}-s)$ is symmetrically distributed around 0,

$$h(b) \leq \mathbb{E}[s] + b + \frac{k}{N-2}(\bar{s}-b)$$

[69] Using that by construction,

$$\Delta_{N-1} = b + \delta\frac{N-2}{N-1}u_{N-2}$$

hypothesis. Hence, since by construction we have that either $\Delta_M = \overline{s} > b$, or

$$\Delta_M = b + \delta\left[\frac{M-1}{N-1}u_{M-1} + \left(1 - \frac{M}{N-1}\right)u_M\right], \tag{34}$$

this implies that $\Delta_M > b$. By induction, the inequality holds for any majority size $M \geq k$.

(b) We thus show that the sequence $(\Delta_M)_{M\geq k}$ is strictly decreasing. We first consider the case where for any $M \geq k$, $\Delta_M < \overline{s}$, and therefore (34) holds, and

$$u_M = h(\Delta_{M+1}) + \Delta_M - h(\Delta_M) - b \tag{35}$$

Suppose by contradiction that $\Delta_{N-1} \geq \Delta_{N-2}$. By the above equations,

$$\begin{aligned}
\Delta_{N-1} &= b + \delta\frac{N-2}{N-1}u_{N-2} \\
&= \left(1 - \delta\frac{N-2}{N-1}\right)b + \delta\frac{N-2}{N-1}\left[h(\Delta_{N-1}) + \Delta_{N-2} - h(\Delta_{N-2})\right] \\
&\leq \left(1 - \delta\frac{N-2}{N-1}\right)b + \delta\frac{N-2}{N-1}\Delta_{N-1}
\end{aligned}$$

where the inequality derives from the strict monotonicity of h_1. Hence $\Delta_{N-1} \leq b$, which contradicts the above result. Therefore $\Delta_{N-1} < \Delta_{N-2}$. We henceforth proceed by induction. Suppose $\Delta_{M'+1} < \Delta_{M'}$ for any $M' \geq M$, and suppose by contradiction that $\Delta_M \geq \Delta_{M-1}$. By (35), using the monotonicity of h_1, we have that

$$u_M < \Delta_M - b, \qquad \text{and} \qquad u_{M-1} \leq \Delta_M - b,$$

and therefore,

$$\Delta_M < \left(1 - \delta\frac{N-2}{N-1}\right)b + \delta\frac{N-2}{N-1}\Delta_M,$$

i.e. $\Delta_M < b$, which is a contradiction. Hence for any $M \geq k$, $\Delta_{M+1} < \Delta_M$, as was to be shown.

We now consider the case where there exists $M \geq k$ such that $\Delta_M = \overline{s}$. This implies that $\Delta_{M+1} < \overline{s}$ as otherwise the explicit expressions of V_M and V_{M+1} would give that

$$\delta\left[\frac{M-1}{N-1}u_{M-1} + \left(1 - \frac{M}{N-1}\right)u_M\right] = 0, \qquad \text{and thus} \qquad \Delta_M = b < \overline{s},$$

which is a contradiction. Hence suppose by contradiction that $\Delta_{N-1} = \overline{s}$, then $\Delta_{N-2} < \overline{s} =$

Δ_{N-1}. Yet the above computations[70] thus yield that $\Delta_{N-1} \leq b < \overline{s}$, which is a contradiction. Therefore, $\Delta_{N-1} < \overline{s}$, and as a consequence, the above computations yield that $\Delta_{N-2} > \Delta_{N-1}$. We again proceed by induction. Suppose $\Delta_{M'+1} < \Delta_{M'}$ for any $M' \geq M$. If $\Delta_M < \overline{s}$, the above computations apply, yielding that $\Delta_M < \Delta_{M-1}$. Hence, suppose by contradiction that $\Delta_M = \overline{s} \geq \Delta_{M-1}$. As noted above, this implies that $\Delta_{M-1} < \overline{s}$ and (35) holds in $M-1$, and thus $u_{M-1} \leq \Delta_M - b$. Moreover, using the explicit expressions of V_{M+1} and V_M,

$$\begin{aligned} u_M &= h(\Delta_{M+1}) - h(\Delta_M) + \delta\Big[\frac{M-1}{N-1}u_{M-1} + \Big(1 - \frac{M}{N-1}\Big)u_M\Big] \\ &< \delta\Big[\frac{M-1}{N-1}u_{M-1} + \Big(1 - \frac{M}{N-1}\Big)u_M\Big] \end{aligned}$$

where the inequality follows from the monotonicity of h. Therefore, $u_M < u_{M-1}$. As a consequence,

$$\begin{aligned} \Delta_M = \overline{s} &\leq b + \delta\Big[\frac{M-1}{N-1}u_{M-1} + \Big(1 - \frac{M}{N-1}\Big)u_M\Big] \\ &< b + \delta\frac{N-2}{N-1}u_{M-1} < \Big(1 - \delta\frac{N-2}{N-1}\Big)b + \delta\frac{N-2}{N-1}\Delta_M, \end{aligned}$$

i.e. $\Delta_M < b < \overline{s}$, which is a contradiction. Hence, for any $M \geq k$, $\Delta_{M+1} < \Delta_M$, as was to be shown.

We then turn to showing that equilibria can be ranked from more to less meritocratic. Consider the class of equilibria characterized by a decreasing decision rule $(\Delta_M)_{M\in\{k,\ldots,N-1\}}$. We refer in the following to an equilibrium by its decision rule $\Delta \equiv (\Delta_M)_{M\in\{k,\ldots,N-1\}}$. Let Δ and Δ' be two equilibria within this class. We now show that

(i) $\Delta_k < \Delta'_k$ implies that $\Delta_M < \Delta'_M$ for any $M \geq k+1$,

(ii) $\Delta_k = \Delta'_k \in [b, \overline{s}]$ implies that $\Delta_M = \Delta'_M < \overline{s}$ for any $M \geq k+1$,

(i) Assume that $\Delta_k < \Delta'_k < \overline{s}$ (computations are analogous in the case $\Delta_k < \Delta'_k = \overline{s}$). By monotonicity, $\Delta_M < \overline{s}$ and $\Delta'_M < \overline{s}$ for any $M \geq k+1$, and thus, with the above notation,

$$\begin{aligned} \Delta_M &= b + \delta\Big[\frac{M-1}{N-1}u_{M-1} + \Big(1 - \frac{M}{N-1}\Big)u_M\Big] \\ &= \Big(1 - \delta\frac{N-2}{N-1}\Big)b + \delta\Big[\frac{M-1}{N-1}[h(\Delta_M) + h_1(\Delta_{M-1})] + \Big(1 - \frac{M}{N-1}\Big)[h(\Delta_{M+1}) + h_1(\Delta_M)]\Big] \end{aligned}$$

[70] Using that as $\Delta_{N-1} = \overline{s}$,

$$\Delta_{N-1} \leq b + \delta\frac{N-2}{N-1}u_{N-2} \leq \Big(1 - \delta\frac{N-2}{N-1}\Big)b + \delta\frac{N-2}{N-1}\Delta_{N-1}$$

Consequently, for any $M \geq k+1$,

$$\begin{aligned}
&h_{2,M}(\Delta_M) - h_{2,M}(\Delta'_M) \\
&\quad = \delta\frac{M-1}{N-1}\Big[h_1(\Delta_{M-1}) - h_1(\Delta'_{M-1})\Big] + \delta\Big(1 - \frac{M}{N-1}\Big)\Big[h(\Delta_{M+1}) - h(\Delta'_{M+1})\Big]
\end{aligned} \tag{36}$$

where the function $h_{2,M}$ is given by

$$h_{2,M}(X) \equiv X - \delta\frac{M-1}{N-1}h(X) - \delta\Big(1 - \frac{M}{N-1}\Big)h_1(X),$$

We note that $h_{2,M}$ is strictly increasing over $[b, \overline{s}]$ [71]. By monotonicity of h_1, we get for $M = k+1$ that

$$h_{2,k+1}(\Delta_{k+1}) - h_{2,k+1}(\Delta'_{k+1}) < \delta\Big(1 - \frac{k+1}{N-1}\Big)\Big[h(\Delta_{k+2}) - h(\Delta'_{k+2})\Big]$$

Suppose by contradiction that $\Delta_{k+1} \geq \Delta'_{k+1}$. Then by monotonicity, $\Delta_{k+2} \geq \Delta'_{k+2}$. By summing Equation (36) in $k+1$ and $k+2$ and rearranging, we get that

$$\begin{aligned}
&\Big[h_{2,k+1}(\Delta_{k+1}) - \delta\frac{k+1}{N-1}h_1(\Delta_{k+1})\Big] - \Big[h_{2,k+1}(\Delta'_{k+1}) - \delta\frac{k+1}{N-1}h_1(\Delta'_{k+1})\Big] \\
&\qquad + \Big[h_{2,k+2}(\Delta_{k+2}) - \delta\frac{k-2}{N-1}h(\Delta_{k+2})\Big] - \Big[h_{2,k+2}(\Delta'_{k+2}) - \delta\frac{k-2}{N-1}h(\Delta'_{k+2})\Big] \\
&= \delta\frac{k}{N-1}\Big[h_1(\Delta_k) - h_1(\Delta'_k)\Big] + \delta\Big(1 - \frac{k+2}{N-1}\Big)\Big[h(\Delta_{k+3}) - h(\Delta'_{k+3})\Big]
\end{aligned}$$

Since for any $M \geq k+1$, the functions $h_{2,M} - \delta\frac{M}{N-1}h_1$ and $h_{2,M} - \delta\frac{N-M}{N-1}h$ are strictly increasing over $[b, \overline{s}]$, the above equality implies that $\Delta_{k+3} \geq \Delta'_{k+3}$. We now proceed by induction: suppose that $\Delta_j \geq \Delta'_j$ for any $j \in \{k+1, ..., M\}$. Then by summing Equation (36) over the indices $k+1, ..., M$ and rearranging,

$$\begin{aligned}
&\Big[h_{2,k+1}(\Delta_{k+1}) - \delta\frac{k}{N-1}h_1(\Delta_{k+1})\Big] - \Big[h_{2,k+1}(\Delta'_{k+1}) - \delta\frac{k}{N-1}h_1(\Delta'_{k+1})\Big] \\
&\qquad + \Big[h_{2,M}(\Delta_M) - \delta\frac{N-M}{N-1}h(\Delta_M)\Big] - \Big[h_{2,M}(\Delta'_M) - \delta\frac{N-M}{N-1}h(\Delta'_M)\Big] \\
&\qquad + \sum_{j=k+2}^{M-1}\Bigg(\Big[h_{2,j}(\Delta_j) - \delta\frac{j}{N-1}h_1(\Delta_j) - \delta\frac{N-j}{N-1}h(\Delta_j)\Big] \\
&\qquad\qquad - \Big[h_{2,j}(\Delta'_j) - \delta\frac{j}{N-1}h_1(\Delta'_j) - \delta\frac{N-j}{N-1}h(\Delta'_j)\Big]\Bigg) \\
&= \delta\frac{k-1}{N-1}\Big[h_1(\Delta_k) - h_1(\Delta'_k)\Big] + \delta\Big(1 - \frac{M}{N-1}\Big)\Big[h(\Delta_{M+1}) - h(\Delta'_{M+1})\Big]
\end{aligned}$$

---

[71] Indeed, we may rewrite the function $h_{2,M}$ as: $h_{2,M}(X) = \Big[1 - \delta\Big(1 - \frac{M}{N-1}\Big)\Big]h_1(X) + \Big[1 - \delta\frac{M-1}{N-1}\Big]h(X)$.

Since for any $j \geq k+1$, the functions $h_{2,j} - \delta \frac{j}{N-1} h_1 - \delta \frac{N-j}{N-1} h$ are strictly increasing over $[b, \overline{s}]$, we get that $\Delta_{M+1} \geq \Delta'_{M+1}$. Hence by induction, we have that $\Delta_M \geq \Delta'_M$ for any $M \geq k+1$. But by summing (36) over all these indices and rearranging yields

$$\begin{aligned}
0 \leq & \Big[h_{2,k+1}(\Delta_{k+1}) - \delta \frac{k}{N-1} h_1(\Delta_{k+1})\Big] - \Big[h_{2,k+1}(\Delta'_{k+1}) - \delta \frac{k}{N-1} h_1(\Delta'_{k+1})\Big] \\
& + \Big[h_{2,N-1}(\Delta_{N-1}) - \delta \frac{1}{N-1} h(\Delta_{N-1})\Big] - \Big[h_{2,N-1}(\Delta'_{N-1}) - \delta \frac{1}{N-1} h(\Delta'_{N-1})\Big] \\
& + \sum_{j=k+2}^{N-2} \Bigg(\Big[h_{2,M}(\Delta_M) - \delta \frac{M}{N-1} h_1(\Delta_M) - \delta \frac{N-M}{N-1} h(\Delta_M)\Big] \\
& \qquad - \Big[h_{2,M}(\Delta'_M) - \delta \frac{M}{N-1} h_1(\Delta'_M) - \delta \frac{N-M}{N-1} h(\Delta'_M)\Big]\Bigg) \\
= & \, \delta \frac{k-1}{N-1} \Big[h_1(\Delta_k) - h_1(\Delta'_k)\Big] < 0
\end{aligned}$$

which is a contradiction. Therefore, $\Delta_{k+1} < \Delta'_{k+1}$. The result then obtains by induction, supposing by contradiction that $\Delta_j < \Delta'_j$ for any $j \in \{k, ..., M-1\}$ and that $\Delta_M \geq \Delta'_M$, and considering the sums of (36) over appropriate indices so as to reach a contradiction.
(ii) We note that the above argument yields that if $\Delta_k = \Delta'_k \in [b, \overline{s}]$, then $\Delta_M = \Delta'_M$ for any $M \geq k+1$. As a consequence, any two distinct equilibria with a decreasing decision rule satisfy either "$\Delta_M < \Delta'_M$ for all $M \geq k$", or "$\Delta_M > \Delta'_M$ for all $M \geq k$".

Lastly, we turn to comparing the equilibria in terms of non-ergodic welfare. Consider two equilibria described by a decreasing decision rule denoted respectively by Δ and Δ' such that $\Delta \prec \Delta'$, and let $(V_i)_{i \in \{1,...,N-1\}}$ and $(V'_i)_{i \in \{1,...,N-1\}}$ be the corresponding equilibrium value functions. For any $M \geq k$, we have by construction that

$$\begin{aligned}
V_M = & \, \mathbb{E}[(s+b)\mathbf{1}\{\hat{s} - s \leq \Delta_M\}] + \mathbb{E}[\hat{s}\mathbf{1}\{\hat{s} - s > \Delta_M\}] \\
& + \delta \mathbb{P}(\hat{s} - s \leq \Delta_M) \Big[\frac{M}{N-1} V_M + \Big(1 - \frac{M}{N-1}\Big) V_{M+1}\Big] \\
& + \delta (1 - \mathbb{P}(\hat{s} - s \leq \Delta_M)) \Big[\frac{M-1}{N-1} V_{M-1} + \Big(1 - \frac{M-1}{N-1}\Big) V_M\Big]
\end{aligned}$$

We first note that $\Delta_k < \Delta'_k$ implies that $\Delta_k < \overline{s}$. Hence, for any $M \geq k$,

$$\Delta_M = b + \delta \Big[\frac{M-1}{N-1} u_{M-1} + \frac{N-M-1}{N-1} u_M\Big],$$

and therefore, for any $M \geq k$,

$$\left[1-\delta\left(1-\frac{M-1}{N-1}\right)[1-\mathbb{P}(\hat{s}-s\leq\Delta'_M)]-\delta\frac{M}{N-1}\mathbb{P}(\hat{s}-s\leq\Delta'_M)\right)\right](V_M-V'_M)$$
$$=\mathbb{E}[(\hat{s}-s-\Delta_M)\mathbf{1}\{\Delta_M<\hat{s}-s\leq\Delta'_M\}]+\delta\mathbb{P}(\hat{s}-s\leq\Delta'_M)\left(1-\frac{M}{N-1}\right)(V_{M+1}-V'_{M+1})$$
$$+\delta(1-\mathbb{P}(\hat{s}-s\leq\Delta'_M))\frac{M-1}{N-1}(V_{M-1}-V'_{M-1}) \qquad (37)$$

Two cases arise depending on whether $\Delta'_k = \overline{s}$. If so, then the result for majority members follows by the usual argument (by contradiction and by induction). Hence, for any $\delta \in [0,(N-1)/N)$, any "meritocratic" equilibrium (i.e. with $\Delta_k < \overline{s}$) is preferred at any majority size by all majority members to the entrenched equilibrium ($\Delta'_k = \overline{s}$).

If $\Delta'_k < \overline{s}$, we need to adapt the arguments in the proof of Lemma 1 and Proposition 2. Suppose by contradiction that $V_{N-1} \leq V'_{N-1}$. Then equation (37) implies that $V_{N-2}-V'_{N-2} \leq V_{N-1}-V'_{N-1} \leq 0$, and thus by induction that $V_{k-1}-V'_{k-1} \leq V_k - V'_k \leq V_{k+1}-V'_{k+1} \leq ... \leq V_{N-1}-V'_{N-1} \leq 0$. However, since $\Delta_k < \Delta'_k < \overline{s}$, we have that

$$b+\delta\frac{k-1}{N-1}(V_{k+1}-V_{k-1}) < b+\delta\frac{k-1}{N-1}(V'_{k+1}-V'_{k-1}),$$

and thus, $V_{k-1}-V'_{k-1} > V_{k+1}-V'_{k+1}$, which contradicts the above inequality. Hence, $V_{N-1} \geq V'_{N-1}$, and (37) implies by induction that $V_{k-1}-V'_{k-1} \geq V_k - V'_k \geq ... \geq V_{N-1}-V'_{N-1} \geq 0$. Therefore, a more meritocratic equilibrium is preferred at any majority size by all majority members to a less meritocratic equilibrium.

Similarly, for any $i \leq k-1$, we have by construction that

$$V_i = \mathbb{E}[s\mathbf{1}\{\hat{s}-s\leq\Delta_{N-1-i}\}]+\mathbb{E}[(\hat{s}+b)\mathbf{1}\{\hat{s}-s>\Delta_{N-1-i}\}]$$
$$+\delta\mathbb{P}(\hat{s}-s\leq\Delta_{N-1-i})\left[\frac{i-1}{N-1}V_{i-1}+\left(1-\frac{i-1}{N-1}\right)V_i\right]$$
$$+\delta(1-\mathbb{P}(\hat{s}-s\leq\Delta_{N-1-i}))\left[\frac{i}{N-1}V_i+\left(1-\frac{i}{N-1}\right)V_{i+1}\right]$$

Hence, for any $i \leq k-1$,

$$\Big[1-\delta\Big(1-\frac{i-1}{N-1}\Big)\mathbb{P}(\hat{s}-s\leq\Delta'_{N-1-i})-\delta\frac{i}{N-1}[1-\mathbb{P}(\hat{s}-s\leq\Delta'_{N-1-i})]\Big](V_i-V'_i)$$
$$=\mathbb{E}\Big[\Big[\hat{s}-s+b+\delta\Big(\frac{i-1}{N-1}u_{i-1}+\frac{N-1-i}{N-1}u_i\Big)\Big]\mathbf{1}\{\Delta_{N-1-i}<\hat{s}-s\leq\Delta'_{N-1-i}\}\Big]$$
$$+\delta\mathbb{P}(\hat{s}-s\leq\Delta'_{N-1-i})\frac{i-1}{N-1}(V_{i-1}-V'_{i-1})$$
$$+\delta(1-\mathbb{P}(\hat{s}-s\leq\Delta'_{N-1-i}))\Big(1-\frac{i}{N-1}\Big)(V_{i+1}-V'_{i+1}) \tag{38}$$

Hence, for δ close to 0, the expectation term on the RHS of (38) is strictly positive. Suppose by contradiction that $V_1 \leq V'_1$. Then, by induction, equation (38) yields that $V_k - V'_k \leq ... \leq V_1 - V'_1 \leq 0$. However, we know from above that $V_{k-1} - V'_{k-1} \geq 0$, hence a contradiction.

Therefore, $V_1 > V'_1$. Working in a similar fashion – by contradiction and by induction using (38) – yields that, for δ small, $V_i > V'_i$ for all $i \in \{1, ..., k-2\}$.

L Proofs of Proposition 8

L.1 Proof of the existence claims in Proposition 8

The properties of the value functions of the two canonical equilibria with homogamic evaluation capability depend on whether $x^\dagger \leq 1/2$. If $x^\dagger \leq 1/2$, they exhibit the same features – monotonicity and concavity/convexity – as their perfect-information counterparts (indeed, the proof of Lemma 1 goes through replacing x by $x^\dagger$). By contrast, if $x^\dagger > 1/2$, the value function in the meritocratic equilibrium (if it exists) now decreases with group size $i \in \{1, ..., N-1\}$ [This observation immediately gives that for $x^\dagger > 1/2$, the meritocratic equilibrium exists for any $s^\dagger > b$.], and is concave for the minority ($i \leq k-1$) and convex for the majority ($i \geq k$). Similarly, in the entrenched equilibrium (if it exists), the value function increases less over $\{k, ..., N-1\}$ than it decreases over $\{1, ..., k-1\}$, whereas with $x^\dagger \leq 1/2$, the opposite holds: the distinction stems from the fact that the (weighted) sum of differences $V^{\text{e}}_{i+1} - V^{\text{e}}_i$ is equal to $(1-2x^\dagger)b$. As a consequence, with $x^\dagger \geq 1/2$, in the entrenchment equilibrium, it is not the case in general that $V^{\text{e}}_i \geq V^{\text{e}}_{N-i-1}$ for any $i \geq k$, while in the meritocratic equilibrium, $V^{\text{m}}_i \leq V^{\text{m}}_{N-i-1}$ for any $i \geq k$ (the curse of control in action).

Let the quantities $Y^\dagger$ and $Z^\dagger$ be given by

$$\begin{cases} Y^\dagger \equiv 1 + \delta \dfrac{k-1}{N-1} x^\dagger \displaystyle\sum_{t=0}^{+\infty} \delta^t \Big(\pi^{\mathrm{e}\dagger}_{k+1,k}(t) - \hat{\pi}^{\mathrm{e}\dagger}_{k,k}(t) \Big) \\ Z^\dagger \equiv 1 + \dfrac{k-1}{N-1} \dfrac{\delta}{1-\delta} (1 - 2x^\dagger) + \delta \dfrac{k-1}{N-1} x^\dagger \displaystyle\sum_{t=0}^{+\infty} \delta^t \Big(\pi^{\mathrm{e}\dagger}_{k+1,k}(t) + \hat{\pi}^{\mathrm{e}\dagger}_{k,k}(t) \Big) \end{cases}$$

where the probabilities $\pi^{\mathrm{e}\dagger}_{i,j}(t)$ (resp. $\hat{\pi}^{\mathrm{e}\dagger}_{i,j}(t)$) are taken *(a)* following the entrenched equilibrium strategies described in Proposition 8, and *(b)* from a majority member's perspective (resp. minority member's perspective) with transition parameter $x^\dagger$ instead of x. Define then $\rho^{\mathrm{e}\dagger}$ as

$$\rho^{\mathrm{e}\dagger} \equiv \begin{cases} \dfrac{x^\dagger}{x} \dfrac{Z^\dagger}{Y^\dagger} & \text{if } Y^\dagger > 0 \\ +\infty & \text{otherwise.} \end{cases}$$

The same argument as the one used in the proof of $\rho^{\mathrm{e}} < +\infty$ [72] yields that for any $\delta \in [0, (N-1)/N)$ and $x^\dagger \in [0, 1)$, $\rho^{\mathrm{e}\dagger} < \infty$.

Similarly, let $\rho^{\mathrm{m}\dagger}$ be defined as

$$\rho^{\mathrm{m}\dagger} \equiv \frac{x^\dagger}{x} \left[1 + \frac{k-1}{N-1} (1 - 2x^\dagger) \delta \sum_{t=0}^{+\infty} \delta^t \left[\Big(\sum_{i=k}^{N-1} \pi^{\mathrm{m}\dagger}_{k+1,i}(t) \Big) - \Big(\sum_{i=k}^{N-1} \pi^{\mathrm{m}\dagger}_{k-1,i}(t) \Big) \right] \right]$$

where the probabilities $\pi^{\mathrm{m}\dagger}_{i,j}(t)$ are taken *(a)* following the meritocratic equilibrium strategies described in Proposition 8, and *(b)* from the perspective of a member of the group with initial size i, with transition parameter $x^\dagger$ instead of x. We show that the thresholds $\rho^{\mathrm{m}\dagger}$ and $\rho^{\mathrm{e}\dagger}$ are the homogamic-evaluation-capability counterparts of ρ^{m} and ρ^{e} in the baseline setting.

The proof of Proposition 8 is analogous to that of Proposition 1. As mentioned, when $x^\dagger \leq 1/2$, the value functions in the entrenched and meritocratic equilibria with homogamic evaluation capability exhibit features similar to the ones of their perfect-information counterparts. Namely, the sequence $(V^{\mathrm{e}\dagger}_M)_{M \geq k}$ remains increasing and concave. By contrast, the monotonicity of the sequence $(V^{\mathrm{m}\dagger}_M)_{M \geq k}$ may differ: it is increasing (and concave) if $x^\dagger \leq 1/2$, whereas it is decreasing (and convex) if $x^\dagger > 1/2$. Moreover, in this latter case it may then be that $V^{\mathrm{e}\dagger}_k < V^{\mathrm{e}\dagger}_{k-1}$. Nonetheless, for $x^\dagger > 1/2$, the sequence $(V^{\mathrm{m}\dagger}_M)_{M \geq k}$ being decreasing implies that its differences $(V^{\mathrm{m}\dagger}_{M+1} - V^{\mathrm{m}\dagger}_M)$ are negative and thus recruiting the minority candidate against an untalented majority candidate is optimal (as $s^\dagger > b$): hence, for $x^\dagger > 1/2$, the meritocratic equilibrium exists whenever $s^\dagger > b$. Lastly, in both cases, because of discounting, a talented majority candidate is still preferred to the minority candidate (with unknown talent) at any majority size.

[72] Cf. Section C.3.

We thus consider $x^\dagger \in [0,1]$ henceforth. As noted above, the argument used in step 1 of the proof of Proposition 1 applies to both equilibria[73], thus yielding that (except in the meritocratic equilibrium for $x^\dagger > 1/2$), the most profitable deviation from these candidate equilibria is when the majority is tight and faces an untalented majority candidate together with an unknown-quality minority one. We thus focus on step 2 and consider one-shot deviations in majority size $M = k$ when the majority candidate is untalented.

A (one-shot) deviation in majority size k from the entrenched strategy (defined in Proposition 8), i.e. picking the minority candidate (of unknown talent) instead of the untalented majority candidate, yields a payoff equal to:[74]

$$\Delta^{\mathrm{e},\dagger} \equiv s^\dagger - b + \delta\frac{k-1}{N-1}xs\sum_{t=0}^{+\infty}\delta^t\Big(\pi^{\mathrm{e}\dagger}_{k+1,k}(t) - \hat{\pi}^{\mathrm{e}\dagger}_{k,k}(t)\Big) + \delta\frac{k-1}{N-1}x^\dagger b\sum_{t=0}^{+\infty}\delta^t\Big(\sum_{i\geq k+1}\hat{\pi}^{\mathrm{e}\dagger}_{k,i}(t)\Big)$$
$$- \delta\frac{k-1}{N-1}x^\dagger b\sum_{t=0}^{+\infty}\delta^t\pi^{\mathrm{e}\dagger}_{k+1,k}(t) - \frac{k-1}{N-1}\frac{\delta}{1-\delta}(1-x^\dagger)b$$

where the probabilities $\pi^{\mathrm{e}\dagger}_{i,j}(t)$ (resp. $\hat{\pi}^{\mathrm{e}\dagger}_{i,j}(t)$) are taken *(a)* following the entrenched equilibrium strategies described in Proposition 8, and *(b)* from a majority member's perspective (resp. minority member's perspective) with transition parameter $x^\dagger$ instead of x. By construction, $s^\dagger/s = x/x^\dagger$. Rearranging yields

$$\Delta^{\mathrm{e},\dagger} = \frac{x}{x^\dagger}s\Big[1 + \delta\frac{k-1}{N-1}x^\dagger\sum_{t=0}^{+\infty}\delta^t\Big(\pi^{\mathrm{e}\dagger}_{k+1,k}(t) - \hat{\pi}^{\mathrm{e}\dagger}_{k,k}(t)\Big)\Big]$$
$$- b\Big[1 + \frac{k-1}{N-1}\frac{\delta}{1-\delta}(1-2x^\dagger) + \delta\frac{k-1}{N-1}x^\dagger\sum_{t=0}^{+\infty}\delta^t\Big(\pi^{\mathrm{e}\dagger}_{k+1,k}(t) + \hat{\pi}^{\mathrm{e}\dagger}_{k,k}(t)\Big)\Big]$$

which yields the result for the existence region of the entrenched equilibrium.

Similarly for the meritocratic equilibrium, consider the (one-shot) deviation of a majority member voting in k the untalented majority candidate instead of the minority one. Such a

[73]For both equilibria when $x^\dagger \leq 1/2$ and for the entrenchment equilibrium when $x^\dagger \geq 1/2$, the argument goes through replacing x by $x^\dagger$ and s by $s^\dagger$ when appropriate. In particular, in the entrenched equilibrium, for $x^\dagger \in [0,1]$, analogous computations yield that

$$\delta\Big(\frac{k-2}{N-1}u^{\mathrm{e}}_{k+1} + \frac{k}{N-1}u^{\mathrm{e}}_k\Big) \leq \frac{\delta\frac{k}{N-1}}{1-\delta\frac{k}{N-1}}\frac{1}{1-\overline{x}}\big(xs - (1-\overline{x})b\big) < s^\dagger - b$$

[74]Indeed, the difference between the expected maximum of both candidates' talents and the expected quality of the majority candidate writes as before $(\overline{x} + (1-\overline{x})x/x^\dagger)s - \overline{x}s = xs$.

deviation yields a payoff equal to:

$$\Delta^{\mathrm{m},\dagger} = b - s^\dagger + \delta\frac{(k-1)}{N-1}(1-x^\dagger)b\sum_{t=0}^{+\infty}\delta^t\Big[\Big(\sum_{i\geq k}\pi^{\mathrm{m}\dagger}_{k+1,i}(t)\Big) - \Big(\sum_{i\geq k}\pi^{\mathrm{m}\dagger}_{k-1,i}(t)\Big)\Big]$$
$$+\,\delta\frac{(k-1)}{N-1}x^\dagger b\sum_{t=0}^{+\infty}\delta^t\Big[\Big(\sum_{i\leq k-1}\pi^{\mathrm{m}\dagger}_{k+1,i}(t)\Big) - \Big(\sum_{i\leq k-1}\pi^{\mathrm{m}\dagger}_{k-1,i}(t)\Big)\Big]$$

i.e. by rearranging,

$$\Delta^{\mathrm{m},\dagger} = -\frac{x}{x^\dagger}s + b\Big[1 + \delta(1-2x^\dagger)\frac{(k-1)}{N-1}\sum_{t=0}^{+\infty}\delta^t\Big[\Big(\sum_{i\geq k}\pi^{\mathrm{m}\dagger}_{k+1,i}(t)\Big) - \Big(\sum_{i\geq k}\pi^{\mathrm{m}\dagger}_{k-1,i}(t)\Big)\Big]\Big]$$

The result for the existence region of the meritocratic equilibrium follows. Lastly, the proof for $\rho^{\mathrm{e},\dagger} < +\infty$ is in Section C.3.

Note moreover that Lemma C.2 holds with the transition probabilities $\pi^{\mathrm{e}\dagger}$ and $\pi^{\mathrm{m}\dagger}$ [75], and this establishes the inequality $\rho^{\mathrm{m}\dagger} < \rho^{\mathrm{e}\dagger}$ for $x^\dagger \leq 1/2$, as well as the inequality $\rho^{\mathrm{m}\dagger} \leq x^\dagger/x$ for $x^\dagger \geq 1/2$ (noted in the text).[76]

L.2 Proof of the welfare claims in Proposition 8

The same argument as the one used in the proof of Proposition (2) yields that, whenever they co-exist, the meritocratic equilibrium is preferred to the entrenchment equilibrium by all members at any majority size.

We now consider ergodic per-period aggregate welfare. We first show that with homogamic evaluation capability, meritocracy dominates entrenchment. To this end, we show that the result of Proposition 3, proved in Online Appendix G, holds replacing x with $x^\dagger \in [0,1]$. Analogous computations to the ones in Online Appendix G show that meritocracy dominates entrenchment if and only if

$$N(N-1)x\left[\frac{x^\dagger}{1-x^\dagger}\frac{k+1}{N} + 1 + \sum_{i=1}^{k-1}\Big(\frac{1-x^\dagger}{x^\dagger}\Big)^i\prod_{l=1}^{i}\frac{k-l}{k+1+l}\right]q\tilde{s}$$
$$> \frac{2x^\dagger}{1-x^\dagger}\left[1 + \sum_{i=1}^{k-1}(i+1)^2\Big(\frac{1-x^\dagger}{x^\dagger}\Big)^i\prod_{l=1}^{i}\frac{k-l}{k+1+l}\right]\tilde{b} \tag{39}$$

[75] Indeed, the proof holds for any $x \in [0,1]$ as the stochastic matrices P and $\hat{P}$ (introduced in the proof of Lemma C.2) remain stochastically monotone and stochastically comparable (with P stochastically dominating $\hat{P}$) for any $x \in [0,1]$.

[76] If $b < s^\dagger$ and $x^\dagger \geq 1/2$, then $\rho^{\mathrm{m}\dagger} \leq x^\dagger/x$, and thus the meritocratic equilibrium exists for all $s/b \geq x^\dagger/x$. Lastly, $s^\dagger$ and $x^\dagger$ both depend on x, and thus the value of $x^\dagger$ constrains the possible values of $s^\dagger$: in particular, for $x^\dagger \geq 1/2$ (and thus $\alpha \leq 1/2$), $s^\dagger$ decreases with $x^\dagger$, and $s^\dagger = 0$ when $x^\dagger = 1$. As a consequence, for any $b > 0$, the inequality $s^\dagger > b$ can only hold for $x^\dagger$ sufficiently below 1.

where $q \geq 1$. By Proposition 8, a necessary condition for meritocracy and entrenchment to exist is $b < s^\dagger$, i.e. $xs > x^\dagger b$. Therefore, a sufficient condition for (39) to be satisfied is

$$N(N-1)\left[\frac{x^\dagger}{1-x^\dagger}\frac{k+1}{N} + 1 + \sum_{i=1}^{k-1}\left(\frac{1-x^\dagger}{x^\dagger}\right)^i \prod_{l=1}^{i}\frac{k-l}{k+1+l}\right]$$
$$> \frac{2}{1-x^\dagger}\left[1 + \sum_{i=1}^{k-1}(i+1)^2\left(\frac{1-x^\dagger}{x^\dagger}\right)^i \prod_{l=1}^{i}\frac{k-l}{k+1+l}\right]$$

By Online Appendix G, the above inequality holds for any $x^\dagger \in [0, 1/2]$, as well as for $x^\dagger$ greater than but close to $1/2$. Moreover, it clearly holds for $x^\dagger$ close to 1. [Numerical simulations suggest it holds for any $x^\dagger[0, 1]$.]

We then turn to the ergodic aggregate welfare comparison of homogamic evaluation capability with respect to perfect information: we show that meritocracy and entrenchment with homogamic evaluation capability are dominated by their perfect-information counterparts. We proceed as in Section 2.2.2.

We first note that in both equilibria, the ergodic distribution of majority sizes with perfect information first-order stochastically dominates the one with homogamic evaluation capability. Using the notation introduced in Section 2.2.2, we denote by $\nu_i^{\mathrm{r}\dagger}$ the ergodic probability of state i at the end of a period in regime $\mathrm{r} \in \{\mathrm{e}, \mathrm{m}\}$, and show that for $\mathrm{r} \in \{\mathrm{e}, \mathrm{m}\}$, the probability distribution $\{\nu_i^{\mathrm{r}}\}$ first-order stochastically dominates $\{\nu_i^{\mathrm{r}\dagger}\}$. Indeed, for $\mathrm{r} \in \{\mathrm{e}, \mathrm{m}\}$, consider the stochastic matrices P^{r} and $P^{\mathrm{r}\dagger}$ associated with the probability distribution over (end-of-period) majority sizes in equilibrium r respectively with perfect information and homogamic evaluation capability, from an outsider's perspective[77]. By construction, both P^{r} and $P^{\mathrm{r}\dagger}$ are stochastically monotone, and the two are stochastically comparable, with $P_{i.}^{\mathrm{r}}$ stochastically dominating $P_{i.}^{\mathrm{r}\dagger}$ for any row index i as $x^\dagger \geq x$. Therefore, the ergodic distribution of majority sizes in equilibrium r with perfect information (first-order) stochastically dominates the one with homogamic evaluation capability.

As a consequence, since the aggregate homophily payoff at a given majority size strictly increases with the majority size, perfect information yields a higher ergodic aggregate homophily payoff than homogamic evaluation capability in equilibrium $\mathrm{r} \in \{\mathrm{e}, \mathrm{m}\}$. Moreover, by Section 2.2.2, the difference in aggregate per-period expected quality between perfect information and homogamic evaluation capability writes as

$$S^{\mathrm{r}} - S^{\mathrm{r}\dagger} = \begin{cases} 0 & \text{if } \mathrm{r} = \mathrm{m}, \\ N(N-1)\left[\nu_{k+1}^{\mathrm{e}\dagger} - \nu_{k+1}^{\mathrm{e}}\right]\dfrac{k+1}{N}x\tilde{s} & \text{if } \mathrm{r} = \mathrm{e}. \end{cases}$$

[77]Namely, for any $i, j \in \{1, ..., k\}$, the matrix component P_{ij}^{r} (resp. $P_{ij}^{\mathrm{r}\dagger}$) is the probability (from an outsider's perspective) that the (end-of-period) majority size moves from $k+i-1$ to $k+j-1$ from one period to another in equilibrium $\mathrm{r} \in \{\mathrm{e}, \mathrm{m}\}$ with perfect information (resp. with homogamic evaluation capability).

Hence, since the probability distribution $\{\nu_i^{\mathrm{e}}\}$ first-order stochastically dominates $\{\nu_i^{\mathrm{e}\dagger}\}$, $S^{\mathrm{r}} - S^{\mathrm{r}\dagger} \geq 0$. Therefore, meritocracy and entrenchment with homogamic evaluation capability are dominated by their perfect-information counterparts in terms of ergodic per-period aggregate welfare.

In order to establish the welfare claim in (i), we show that (perfect-information) entrenchment dominates full-entrenchment. The aggregate ergodic quality in the full-entrenchment equilibrium writes as $S^{\mathrm{f}} = N(N-1)\overline{x}\tilde{s}$, and thus using the computations of Section 2.2.2, the difference between the ergodic efficiency of an entrenched and fully-entrenched organization is given by

$$S^{\mathrm{e}} - S^{\mathrm{f}} = N(N-1)\left[1 - \nu_{k+1}^{\mathrm{e}}\frac{k+1}{N}x\right]\tilde{s}$$

Similarly, the difference ergodic homophily benefits is given by

$$B^{\mathrm{e}} - B^{\mathrm{f}} = \sum_{i=k+1}^{N} \nu_i^{\mathrm{e}}\left[i(i-1) + (N-i)(N-i-1) - N(N-1)\right]$$

Building on Online Appendix F, explicit computations[78] then yield that $q(S^{\mathrm{e}}-S^{\mathrm{f}})+B^{\mathrm{e}}-B^{\mathrm{f}} > 0$ for any $s > b$, hence the result.

The first part of the welfare claim in (ii) stems from the explicit expressions of ρ^{m} and $\rho^{\mathrm{m}\dagger}$ which imply that for δ close to 0,

$$\rho^{\mathrm{m}} = 1 + (1-2x)\frac{k-1}{N-1}\delta + O(\delta^2), \qquad \text{and} \qquad \rho^{\mathrm{m}\dagger} = \frac{x^{\dagger}}{x}\left[1 + (1-2x^{\dagger})\frac{k-1}{N-1}\right]\delta + O(\delta^2),$$

and thus $\rho^{\mathrm{m}} < \rho^{\mathrm{m}\dagger}$. The second part derives from the above results, namely that meritocracy and entrenchment with homogamic evaluation capability are dominated by their perfect-information counterparts, and that meritocracy dominates entrenchment with homogamic evaluation capability as well as with perfect information.

[78] With the explicit expressions for the ergodic probabilities ν_i^{e} derived in Online Appendix F, $q(S^{\mathrm{e}} - S^{\mathrm{f}}) + B^{\mathrm{e}} - B^{\mathrm{f}}$ has the same sign as

$$\left[N(N-1)\left(1 - \frac{k+1}{N}x\right)q\tilde{s} + 2(k+1)(1-k)\tilde{b}\right] + \sum_{i=1}^{k-1}\left(\frac{1-x}{x}\right)^{i}\prod_{l=1}^{i}\frac{k-l}{k+1+l}\left[N(N-1)q\tilde{s} + 2(i+k+1)(i-k+1)\tilde{b}\right]$$

The result obtains by noting that for any $x \leq 1/2$,

$$N(N-1)\left(1 - \frac{k+1}{N}x\right) > 2(k+1)(1-k),$$

and that for any $i \in \{1, ..., k-1\}$, $2(i+k+1)(i-k+1) > 2(k+2)(2-k) > -N(N-1)$.

M Proof of Proposition 9

[We provide in the Supplementary Materials explicit computations for the case $l = k - 2$, which may be helpful to get the spirit of the proof in the general case.]

General proof for existence. Let $s = b > 0$. Consider any $l \in \{1, ..., k-2\}$ and the strategy of super-entrenchment to level l, denoting by V_i the corresponding value function and u_i its first-difference. Since $s = b$, the usual computations[79] (see proof of Lemma 1) yield that for any $i \geq k + l$ and for any $i \leq k - 2 - l$, $u_i = 0$. The above computations then apply, using (**??**) for group sizes $i \in \{k, ..., k + l - 1\}$, (**??**) for group sizes $i \in \{k - 2 - l, ..., k - 2\}$, and (**??**) for group size $k - 1$. The result obtains by continuity for s/b in a neighbourhood of 1.

General proof for uniqueness. We now show that, for s/b close to 1, super-entrenchment at level l is the unique symmetric MPE such that a stronger majority makes more meritocratic recruitments. Hence, we consider the class of equilibria such that a stronger majority makes more meritocratic recruitments, and show that, for any candidate equilibrium within this class, for $s = b > 0$, the majority is super-entrenched in $k + l$. By monotonicity, this implies that all candidate equilibria within this class must feature an entrenched majority at majority sizes $M \in \{k, ..., k + l\}$. We then show that the minority best-replies to this strategy by voting for the in-group candidate whenever it may be pivotal, i.e. at any majority size $M \leq k + l - 1$.

We begin by noting that when $s = b$, a group's flow payoff whenever it is pivotal does not depend on its being meritocratic or entrenched (as the difference between the two writes as $x(s - b) = 0$). Moreover, for $s = b$, the flow differential payoff in the expression of u_i writes as $[\Lambda(i) - \Lambda(i + 1)]b$ (resp. $[\Lambda(i) - \Lambda(i + 1)](1 - 2x)b$) if the minority follows entrenchment (resp. meritocracy) at majority sizes i and $i + 1$, as $[\Lambda(i) - \Lambda(i + 1)]b - 2x\Lambda(i)b$ if the minority follows meritocracy at majority size i and entrenchment at majority size $i + 1$, and as $[\Lambda(i) - \Lambda(i + 1)]b + 2x\Lambda(i + 1)b$ if the minority follows entrenchment at majority size i and meritocracy at majority size $i + 1$. In particular, the flow-payoff term in u_{k+l-1} writes as $\Lambda(k+l-1)b$ if the minority is entrenched at majority size $k+l-1$ (resp. $\Lambda(k+l-1)(1-2x)b$ if it votes meritocratically). By contrast, for any $i \geq k + l$, the flow-payoff term in u_i is equal to 0.

We now show that the majority is always entrenched in $k + l$, i.e. that in any equilibrium,

$$\frac{k + l - 1}{N - 1} u_{k+l-1} + \left(1 - \frac{k + l}{N - 1}\right) u_{k+l} > 0$$

Suppose by contradiction that the majority votes meritocratically at size $k + l$ (i.e. that

[79] This could be seen by using the recursive expressions for the sequence $(u_i)_i$ and supposing by contradiction that $u_i \neq 0$ for some $i \geq k + l$ or $\leq k - 2 - l$.

the above LHS is weakly lower than 0). Suppose first that $u_{k+l} \leq 0$. Hence, the recursive expression of u_i for $i \geq k+l$ is given by (4) and yields[80] that $u_{k+l-1} \leq u_{k+l} \leq ... \leq u_{N-2} \leq 0$. Then, summing up (and rearranging) the recursive expression of u_{k+l-1} and u_i for $i \geq k+l$ (and rearranging) yields on the LHS a (positively) weighted sum of u_i, $i \geq k+l-1$, which is thus negative, and on the RHS the sum of flow-payoff term in u_{k+l-1}, which is strictly positive, and of a term proportional to u_{k+l-2}. Therefore, $u_{k+l-2} < 0$. We proceed by induction in order to show that $u_i < 0$ for any $i \in \{k-1, ..., k+l-2\}$. Let $M \in \{k, ..., k+l-2\}$, and suppose $u_i \leq 0$ for any $i \geq M$. Summing and rearranging as above the recursive expressions of the differential value function u_i over indices $i \in \{M, ..., N-2\}$ gives on the LHS a weighted sum of u_i for $i \in \{M, ..., N-2\}$, which is weakly negative with the induction hypothesis, while on the RHS a first term proportional to u_{M-1} and a second term which is the sum of the flow-differential payoffs, equal either to $\Lambda(M)(1-2x)b$, $\Lambda(M)b$ or $[\Lambda(M)+\Lambda(M+1)2x]b$, which is thus strictly positive. Therefore, $u_{M-1} < 0$.

Hence, by induction, $u_i < 0$ for any $i \in \{k-1, ..., k+l-2\}$. Therefore the majority is meritocratic at any majority size $i \geq k$. As a consequence, the flow differential payoffs in the expression of u_i for $i \leq k-1$ write as $[\Lambda(i)-\Lambda(i+1)](1-2x)b > 0$ for any $i \in \{k-l-1, ..., k-2\}$, and 0 for any $i \leq k-l-2$.

Suppose by contradiction that $u_{k-l-1} \leq 0$. Then the recursive expression of u_i for $i \leq k-l-2$ is given by (6) and yields that $u_{k-l-1} \leq ... \leq u_1 \leq 0$. Furthermore, since the flow differential payoffs are positive for $i \in \{k-l-1, ..., k-2\}$, we have that $u_i \leq 0$ for $i \in \{1, ..., k-1\}$. Therefore the minority votes meritocratically whenever it is pivotal. Hence, the sum of the flow differential payoffs over all indices $i \in \{1, ..., N-2\}$ writes as

$$2\Lambda(k)(1-2x)b + [1-2\Lambda(k)](1-2x)b = (1-2x)b > 0$$

where the second term is the flow differential payoff in u_{k-1}. Yet this contradicts $u_i \leq 0$ for all $i \in \{1, ..., N-2\}$.

Hence $u_{k-l-1} > 0$. The recursive expressions of the differential value function (6) now yield that $0 < u_1 < ... < u_{k-l-1}$. Supposing by contradiction that $u_{k-l} \leq 0$ yields again that $u_i \leq 0$ for $i \in \{k-l, ..., k-1\}$. Hence by summing the recursive expressions of u_i for $i \in \{k-l, ..., N-2\}$ and rearranging yields on the LHS a weighted sum of the differential value function u_i for $i \in \{k-l, ..., N-2\}$, which is weakly negative, while on the RHS, a term proportional to u_{k-l-1} (and thus strictly positive) and the sum of the flow differential payoffs, which is strictly positive. This is a contradiction, and thus $u_{k-l} > 0$. Using repeatedly the same argument, we have by induction that $u_i > 0$ for any $i \leq k-2$, and as a consequence, the

[80] This can be seen by supposing by contradiction that $u_{N-2} > 0$, and reaching a contradiction a contradiction using (4). The result thus obtains by downward induction, using again (4).

minority is entrenched whenever it has size $i \in \{k-l, ..., k-2\}$, i.e. whenever the majority has size $i \in \{k+1, ..., k+l-1\}$. Summing again the recursive expression of the differential value function u_i over indices $i \geq k-1$ yields after rearranging, on the LHS a weighted sum of the differential value function u_i for $i \in \{k-1, ..., N-2\}$, which is weakly negative, while on the RHS, a term proportional to u_{k-2} (and thus strictly positive) and the sum of the flow differential payoffs, which is equal to $[1-\Lambda(k)](1-2x) > 0$. Hence the RHS is strictly positive, which is a contradiction. Therefore, $u_{k+l} > 0$, and thus using the recursive expression of u_i for $i \geq k+l$ (namely (4) as we suppose that the majority votes meritocratically at size $k+l$), we have that $u_{k+l-1} > u_{k+l} > u_{k+l+1} > ... > u_{N-2} > 0$. This establishes that the majority would strictly benefit by deviating to entrenchment when it has size $k+l$ [81], and thus contradicts the assumption that the majority votes meritocratically at size $k+l$.

Hence the majority is entrenched when it has size $k+l$. Note that this implies that $u_{k+l} = u_{k+l+1} = ... = u_{N-2} = 0$. This establishes the uniqueness of the super-entrenchment at level l within the class of equilibria such that a stronger majority makes more meritocratic recruitments. Furthermore, the argument implies that in any symmetric MPE in weakly undominated strategies, the majority is entrenched when it has size $k+l$.

Lastly, as claimed in footnote 36, there may exist non-monotonic symmetric MPEs in weakly undominated strategies. For instance, with $N = 6$ and parameter values such that: $2(1-\Lambda_3-\Lambda_4) < (1-\Lambda_3)4x < 3-5\Lambda_3-\Lambda_4$, the following strategies describe such an equilibrium for any s/b sufficiently close to 1 and δ sufficiently close to 0:

- vote meritocratically whenever one's group has size 3 (tight majority) – i.e. vote for the in-group candidate if and only if he is at least as talented as the out-group candidate, and vote for the out-group candidate otherwise;
- follow entrenchment otherwise – i.e. always vote for the in-group candidate.

N Proof of Proposition 10

We first show that meritocratic equilibrium *strategies* are no longer so when candidates reapply, for s/b in some interval $[1, \rho^{\text{m}} + \epsilon)$ with $\epsilon > 0$. We then show that the meritocratic equilibrium *path* starting from an initial state with empty storage is no longer an equilibrium path for s/b in some interval $[1, \rho^{\text{m}} + \epsilon)$ with $\epsilon > 0$: an equilibrium may be observationally

[81] Indeed, facing a talented minority candidate and an untalented majority one, the differential payoff would be given by

$$\delta\left[\frac{k+l-1}{N-1}u_{k+l-1} + \left(1 - \frac{k+l}{N-1}\right)u_{k+l}\right] > 0$$

equivalent to a meritocratic equilibrium by exhibiting the same recruitment path, without necessarily be meritocratic off the equilibrium path (more on this below).

We define the meritocratic equilibrium as an equilibrium in which the majority always recruits the best candidate available[82] for any stocks of candidates, and look for necessary conditions for the meritocratic equilibrium to exist. We show the latter are more often binding when candidates reapply than when they cannot. Namely, when candidates reapply, we exhibit one deviation that is profitable for s/b a bit above ρ^{m} (and for all $s/b \in [1, \rho^{\mathrm{m}}]$). Note that we do not derive a sufficient condition for existence.

Two effects (which we will successively illustrate) are at play, shrinking the existence region of meritocracy: (i) the ability to recall a talented minority candidate increases the value of entrenchment; and (ii) the preferential treatment given by the majority to its in-group talented candidate(s) in store makes an incumbent majority with a large number of talented minority candidates in store less willing to relinquish control.

To illustrate both forces at play, consider first $x = 1/2$ (so that $\rho^{\mathrm{m}} = 1$), and $s/b = 1$. Suppose the majority has size k, and no talented majority candidate available[83] but an infinite number of talented minority ones in store. Recruiting a talented minority candidate instead of an untalented majority one gives a differential payoff equal to

$$s - b + \delta \frac{k-1}{N-1}\left(\frac{s}{1-\delta} - V_{k+1,0,\infty}\right) = \delta \frac{k-1}{N-1}\left(\frac{s}{1-\delta} - V_{k+1,0,\infty}\right)$$

where $V_{k+1,0,\infty}$ is the majority value function when it has size $k+1$, no talented majority candidate in store and an infinite number of talented minority ones in store. Since for $x = 1/2$, a majority with size $k+1$ can secure in each period an (expected) flow quality payoff equal to $\tilde{s}$, and for at least the first two periods, an (expected) flow homophily payoff equal to $\tilde{b}/2$ [84], we have that $V_{k+1,0,\infty} > s/(1-\delta)$. Furthermore, as the majority cannot do better than $\tilde{s}$ in terms of flow quality payoff, the term $[s/(1-\delta) - V_{k+1,0,\infty}]$ does not decrease with s, but strictly decreases with b. Therefore, the above differential payoff is strictly negative for any s/b in an upper neighbourhood of 1. Because of time discounting ($\delta_0 < 1$), the result holds when the majority has in store a sufficiently large finite number of talented minority candidates. Hence, for $x = 1/2$, there exists a strictly profitable deviation away from meritocracy for $s/b \in [\rho^{\mathrm{m}}, \rho^{\mathrm{m}} + \epsilon)$.

Consider now $x < 1/2$ (so that $\rho^{\mathrm{m}} > 1$), and $s/b = \rho^{\mathrm{m}}$. A necessary condition for the meritocratic equilibrium to exist is that a repeated deviation towards entrenchment whenever

[82] Namely the best candidate among current-period and stored candidates, breaking ties in favor of in-group candidates as before.

[83] Namely, it has no such candidate in store, and the current-period majority candidate is untalented.

[84] In particular, reverting to the meritocratic strategy yields to the current majority group an (expected) flow payoff equal to $\tilde{s} + \tilde{b}/2$ as long as it retains control over the organization, and equal to $\tilde{s}$ after it has relinquished it to the other group.

the majority is tight ($M = k$) and has no talented majority candidate available and exactly one talented minority candidate available, be non profitable. Upon permanently deviating to entrenchment, the majority has one talented minority candidate in store, and either size k or $k+1$. Yet, for $x < 1/2$, an *entrenched* majority's value function strictly increases with the number of talented minority candidates in store[85]. Hence, when candidates reapply, a permanent deviation away from meritocracy becomes more profitable. Furthermore, an inspection of the additional payoff due to storability shows that the latter increases with s and decreases with b. Intuitively, this derives from the fact that having a talented minority candidate in store leads to the latter being recruited (at some point, with strictly positive probability) instead of a (talented or untalented) in-group candidate or an untalented out-group candidate, thus yielding a positive quality gain and a positive homophily loss with respect to the payoff when candidates cannot reapply. Therefore, since in the absence of storability, we have the equivalence between the profitability of one-shot and permanent deviations[86], there exists a profitable deviation away from meritocracy for $s/b > \rho^{\mathrm{m}}$ (and for all $s/b \in [1, \rho^{\mathrm{m}}]$) , i.e. the existence region of meritocracy shrinks.

Finally we show that the meritocratic equilibrium *path* starting from an initial state with empty storage is no longer an equilibrium path for s/b in some interval $[\rho^{\mathrm{m}}, \rho^{\mathrm{m}} + \epsilon)$ with $\epsilon > 0$. We first note that, on the meritocratic equilibrium path starting from an initial state with empty storage, storage is never used[87]. Considering the repeated deviation to entrenchment described above yields that, for $x < 1/2$, there exists a strictly profitable deviation away from this equilibrium path for s/b slightly above ρ^{m} (and for all $s/b \in [1, \rho^{\mathrm{m}}]$). Hence, when $x < 1/2$, then for s/b in some interval $[\rho^{\mathrm{m}}, \rho^{\mathrm{m}} + \epsilon)$ with $\epsilon > 0$, the meritocratic equilibrium path starting from an initial state with empty storage is no longer so.

[85]Indeed, an entrenched majority solves an optimal control problem. Moreover, as $x < 1/2$, the majority faces two untalented current-period candidates with a strictly positive probability ($1 - 2x > 0$), in which case, whenever it is not tight ($M > k$) and whenever it has a talented minority candidate in store, it recruits the latter, thus receiving a strictly positive differential payoff with respect to the empty-storage state. Indeed, the differential payoff from recruiting a stored talented minority candidate instead of an untalented majority candidate whenever the majority is not tight, is bounded below by:

$$s - b - x(s-b)\frac{\delta k/(N-1)}{1 - \delta k/(N-1)} > (1-x)(s-b) > 0$$

[86]Hence, when candidates cannot reapply, the above repeated deviation yields a zero differential payoff for $s/b = \rho^{\mathrm{m}}$.

[87]Indeed, as we assume $\alpha = 0$, the organization faces at most one new talented candidate each period, and on the meritocratic equilibrium path, recruits her/him.

O Hierarchies and the glass ceiling

For simplicity, we look at the continuous-time version of our model. Consider a large two-tier organization with a mass 1 of senior positions and a mass $J > 1$ of junior positions. A higher J corresponds to a "more pyramidal" organization. Between times t and $t + dt$, a fraction $\chi^S dt$ of seniors departs and is replaced by juniors promoted to seniority; a fraction $\chi^J J dt$ of juniors departs as well. To offset these two flows out of the junior pool, a fraction $\hat{\chi} J dt$ of new juniors is recruited (where $J\hat{\chi} = \chi^S + J\chi^J$). The flow of talented majority (minority) candidates is $X dt$. We will assume that $X \leq J\hat{\chi}$ (otherwise the organization would be homogenous, and the absence of minority juniors would deprive us of an analysis of the glass ceiling). Seniors have control over hiring and promotion decisions.

As noted in the text, a glass ceiling in such hierarchical organizations results from control being located at the senior level. This operates through two channels:

- *Concern for control:* as earlier in the paper, control allows groups to engage in favoritism. Because control is located at the senior level, this in turn implies some discrimination in promotions, which in general exceeds that at the hiring level (if any). A concern for control and the concomitant discrimination may arise even in large organizations, either because of shocks, or because the talent pool is larger in the minority.

- *Differential mingling effect:* for organizational reasons, senior members tend to hang around more with senior members than with junior ones. Their homophily concerns are therefore higher for promotions than for hiring decisions.

Because the second effect is at this stage of the paper newer, we illustrate it through a simple example, which can be much enriched in ways that we later discuss. Assume that senior members enjoy (expected lifetime) homophily benefits from in-group senior and junior members, which we denote respectively by b^S and b^J. The differential mingling effect is captured by $b^S > b^J$. A fraction $x \leq 1/2$ of new hires are in-group talented juniors, and similarly for the out-group ones: $xJ\hat{\chi}dt = Xdt$. Talent is observed prior to hiring. A talented member brings quality benefits to seniors equal to s^J when junior, and $s^S > s^J$ when senior. Assume that $s^l > b^l$ at both levels $l \in \{J, S\}$, and that $s^S - s^J > b^S - b^J$ (these two conditions generalize the previous assumption that quality matters to the majority).

In this framework, majority members are never worried about losing control, as the promotion of those who wil bring them the highest net benefits will always be tilted in favor of in-group juniors. This leads us to focus on the *majority's pecking order*: A promotion yields discounted net benefit to a majority senior member equal to 1) $s^S - s^J + b^S - b^J$ in the case of an in-group talented member; 2) $s^S - s^J$ for an out-group talented member; 3) $b^S - b^J$ for an in-group untalented member; 4) 0 for an out-group untalented member. This pecking order implies that promotion decisions will be tilted in favor of in-group members (except in the non-generic case in which all talented juniors are promoted and no untalented one is). In contrast, the junior population is balanced in composition; indeed, there is no rationale for the majority to discriminate at the hiring state as long as $s^J > b^J$.

When $X < \chi^S < 2X$, i.e. equivalently $x < 1/[1 + J\chi^J/\chi^S] < 2x$, in steady state the organization promotes all talented in-group juniors, a fraction z of talented out-group juniors, and no untalented juniors. The flows in and out of the junior and senior pools must balance, yielding respectively: $J\hat{\chi} = \chi^S + J\chi^J$, and $J\hat{\chi}x(1+z) = \chi^S$.

We define the glass ceiling index as the relative probability of promotion of talented majority and minority members, minus 1:[88]

$$\gamma \equiv \frac{1}{z} - 1 = \frac{2X - \chi^S}{\chi^S - X} \in (0, \infty)$$

In this region, the glass ceiling index is invariant with how pyramidal the organization is (J)[89], decreases with the frequency of senior-level vacancies (χ^S) and increases with the flow of talented candidates (X). Covering all parameter regions, the glass ceiling index is monotonic with χ^S/X.[90]

Proposition O.1. ***(Glass ceiling)*** *In the hierarchical organization's steady state, hiring at the junior level is meritocratic. By contrast, there exists a glass ceiling for minority juniors.*

This environment can be enriched in interesting ways. First, one may distinguish between talent and "senior potential"; only a fraction of talented members have the potential to make a more important contribution at the senior level; furthermore it may take time for the organization to discover who has such senior potential (there is a time of reckoning). Second, talented members may have outside opportunities. Talented women may then quit the organization due to a discouragement effect: either they have been identified as lacking senior potential (their male counterparts by contrast staying in the organization), or the delay in being promoted is not worth the wait. Finally, the possibility of outside recruitment at the senior level would impact the glass ceiling effect.

[88] This definition of the glass ceiling index only looks at flows and is a conservative estimate of the glass ceiling; indeed, were we to look at stock, the glass ceiling effect would be stronger because the share of talented minority juniors promoted to seniority (over the whole stock of such juniors) would be below z (whenever $z < x$, the steady state of the junior pool features a mixture of talented minority and untalented majority juniors).

[89] An increase in J has two opposite effects: it makes it more difficult for a junior to be promoted, and talented minority members are the first to be left out; but it also makes talented juniors scarcer in the junior pool, increasing the minority members' probability of promotion.

[90] Indeed, for $\chi^S > 2X$, the senior majority hires all talented juniors and (some) untalented in-group juniors, and thus $\gamma = 0$, whereas for $\chi^S < X$, it promotes no out-group talented juniors, only talented in-group ones, and thus we set $\gamma = +\infty$.

P Negative homophily

As claimed in the text (see footnote 3), the case $\tilde{b} < 0$, corresponding to *negative homophily*, can be accommodated in our model. Indeed, the set of possible flow payoffs in any period writes as $\{\tilde{s}, 0, \tilde{s} + \tilde{b}, \tilde{b}\}$. Hence, for $\tilde{b} < 0$, two cases must be distinguished:

- $\tilde{s} + \tilde{b} < 0$ (i.e. $-1 < \tilde{s}/\tilde{b} < 0$): the majority always votes for the minority candidate. The majority size converges to k, which is an absorbing state. The majority then switches and control alternates between the two groups.
- $\tilde{s} + \tilde{b} > 0$ (i.e. $\tilde{s}/\tilde{b} < -1$): there always exists an equilibrium in which the majority votes for the most talented candidate with a tie-breaking rule in favor of the minority candidate.[91]

[91] Indeed, the same computations as in the proof of Lemma 1 (see Online Appendix B) yield that, letting $u_i \equiv V_{i+1} - V_i$, with V_i the value function with such strategies, $0 < u_1 < ... < u_{k-1}$ and $u_{k-1} > ... > u_{N-2} > 0$, with

$$u_{k-1} = -(1-2x)b + \delta x \left[\frac{k-1}{N-1} u_k + u_{k-1} + \frac{k-2}{N-1} u_{k-2} \right],$$

and thus in particular,

$$\left[1 - 2\delta x \frac{N-2}{N-1} \right] u_{k-1} < -(1-2x)b.$$

As a consequence, deviations that yield a lower current-period flow payoff, together with a lower (in a first-order stochastic sense) distribution of next-period in-group sizes are strictly unprofitable. Moreover, as $0 < u_{N-2} < ... < u_{k-1}$, the deviation differential payoff for the majority from picking its in-group candidate instead of an at-least-as-talented out-group candidate (hence opting for a higher distribution of next-period in-group sizes at the expense of a lower current-period flow payoff) is maximal when both candidates have the same talent and the majority has size k. It then writes as

$$b + \delta \frac{k-1}{N-1} (u_{k-1} + u_k) < b + \delta \frac{N-2}{N-1} u_{k-1} < 0$$

using the above upper bound on u_{k-1}. Therefore, such a deviation is never profitable for the majority.